```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         September 25, 2001, 14:42:05 ; Search time 20.45 Seconds
                                                (without alignments)
                                                1746.089 Million cell updates/sec

Title:          US-09-328-877A-8
Perfect score:  589
Sequence:       1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:  OLIGO
                Gapop 60.0 , Gapext 60.0

Searched:       412676 seqs, 60623988 residues

Word size :     0

Total number of hits satisfying chosen parameters:      412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :      A_Geneseq_0601:*
                1:  /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                      Description
-----------------------------------------------------------------------
    1     589   100.0    589  20  AAY22466                Human mammary sel-
    2     589   100.0    589  22  AAB59198                Human mammary sel-
    3     559    94.9    559  20  AAY22467                Human mammary sel-
    4     559    94.9    559  22  AAB59199                Human mammary sel-
    5     557    94.6    589  21  AAB01204                Human GTPase assoc
    6     540    91.7    540  20  AAY22465                Human hippocampal
    7     540    91.7    540  20  AAY22468                Human mammary sel-
    8     540    91.7    540  22  AAB59197                Human hippocampal
    9     540    91.7    540  22  AAB59200                Human mammary sel-
   10     540    91.7    545  20  AAY22464                Human hippocampal
   11     540    91.7    545  22  AAB59196                Human hippocampal
```

```
   12     540    91.7    553  20  AAY22463                Human hippocampal
   13     540    91.7    553  22  AAB59195                Human hippocampal
   14     540    91.7    592  20  AAY22462                Human hippocampal
   15     540    91.7    592  22  AAB59194                Human hippocampal
   16     540    91.7    626  20  AAY22469                Human 6myc-N-sel-1
   17     540    91.7    626  22  AAB59201                Protein encoded by
   18     540    91.7    627  20  AAY22461                Human hippocampal
   19     540    91.7    627  22  AAB59193                Human hippocampal
   20     540    91.7    666  20  AAY22471                Human C-term mychi
   21     540    91.7    666  22  AAB59203                C-terminal mychis
   22     540    91.7    669  20  AAY22470                Human Cterm V5 his
   23     540    91.7    669  22  AAB59202                C-terminal V5 His
   24      35     5.9    122  21  AAG01075                Human secreted pro
   25      12     2.0    587  20  AAY03204                Amino acid sequenc
   26       9     1.5    257  21  AAB27862                Protein fragment e
   27       9     1.5    472  18  AAW23598                Human LYST2 polype
   28       9     1.5    472  20  AAY32120                Human LYST-2 prote
   29       9     1.5    703  18  AAW23599                Mouse LYST2 polype
   30       9     1.5    789  20  AAY32131                Human LYST-2 prote
   31       8     1.4     13  17  AAR91634                Alpha chemokine GR
   32       8     1.4     14  17  AAR91631                Alpha chemokine GR
   33       8     1.4     14  17  AAR91632                Alpha chemokine GR
   34       8     1.4     15  17  AAR91623                Alpha chemokine GR
   35       8     1.4     16  17  AAR91620                Alpha chemokine GR
   36       8     1.4     16  17  AAR91621                Alpha chemokine GR
   37       8     1.4     73  14  AAR36771                MIP-2alpha.  Homo
   38       8     1.4     73  14  AAR36772                MIP-2beta.  Homo s
   39       8     1.4     73  16  AAR66699                Human gro-beta che
   40       8     1.4     73  16  AAR66700                Human gro-gamma ch
   41       8     1.4     73  16  AAR66698                Human gro-alpha ch
   42       8     1.4     73  17  AAR93194                Protein used to ge
   43       8     1.4     73  18  AAW18024                Human chemokine gr
   44       8     1.4     73  18  AAW18025                Human chemokine gr
   45       8     1.4     73  18  AAW18026                Human chemokine gr


                              ALIGNMENTS


RESULT   1
AAY22466
ID   AAY22466 standard; Protein; 589 AA.
XX
AC   AAY22466;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human mammary sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;
KW   mammary gland; therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99702.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 60-63; 91pp; English.
XX
```

```
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human mammary sel-10 protein. The polypeptides
CC   can be used to alter presenilin function. Compounds which inhibit either
CC   the expression or the activity of the human sel-10 polypeptides may
CC   reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and
CC   therefore may be useful for the prevention or treatment of Alzheimer's
CC   disease.
XX
SQ   Sequence   589 AA;

  Query Match             100.0%;  Score 589;  DB 20;  Length 589;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  589;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MSKPGKPTLNHGLVPVDLKSAKEPLPHQTVMKIFSISIIAQGLPFCRRRMKRKLDHGSEV 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 mskpgkptlnhglvpvdlksakeplphqtvmkifsisiiaqglpfcrrrmkrkldhgsev 60

Qy     61 RSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 rsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrritsvqpptglqewl 120

Qy    121 KMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEP 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 kmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpkelalyvlsflep 180

Qy    181 KDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFIHSPWKSAYIRQ 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 kdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfihspwksayirq 240

Qy    241 HRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLV 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 hridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvwsavtgkclrtlv 300

Qy    301 GHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcmhlhekrvvsgs 360

Qy    361 RDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 rdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwdpetetclhtlq 420

Qy    421 GHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGN 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgmelkdnilvsgn 480

Qy    481 ADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIR 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 adstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvklwdlktgefir 540

Qy    541 NLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 nlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 589

RESULT   2
AAB59198
ID   AAB59198 standard; protein; 589 AA.
XX
AC   AAB59198;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human mammary sel-10-1 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 79-82; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   589 AA;

  Query Match             100.0%;  Score 589;  DB 22;  Length 589;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  589;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MSKPGKPTLNHGLVPVDLKSAKEPLPHQTVMKIFSISIIAQGLPFCRRRMKRKLDHGSEV 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 mskpgkptlnhglvpvdlksakeplphqtvmkifsisiiaqglpfcrrrmkrkldhgsev 60

Qy     61 RSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 rsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrritsvqpptglqewl 120

Qy    121 KMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEP 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 kmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpkelalyvlsflep 180

Qy    181 KDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFIHSPWKSAYIRQ 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 kdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfihspwksayirq 240

Qy    241 HRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLV 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 hridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvwsavtgkclrtlv 300

Qy    301 GHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcmhlhekrvvsgs 360

Qy    361 RDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 rdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwdpetetclhtlq 420

Qy    421 GHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGN 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgmelkdnilvsgn 480

Qy    481 ADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIR 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 adstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvklwdlktgefir 540

Qy    541 NLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
          |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    541 nlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 589


RESULT   3
AAY22467
ID   AAY22467 standard; Protein; 559 AA.
XX
AC   AAY22467;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human mammary sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;
KW   mammary gland; therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99702.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 63-66; 91pp; English.
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human mammary sel-10 protein. The polypeptides
CC   can be used to alter presenilin function. Compounds which inhibit either
CC   the expression or the activity of the human sel-10 polypeptides may
CC   reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and
CC   therefore may be useful for the prevention or treatment of Alzheimer's
CC   disease.
XX
SQ   Sequence   559 AA;


  Query Match             94.9%;  Score 559;  DB 20;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  559;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     31 MKIFSISIIAQGLPFCRRRMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTT 90
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 mkifsisiiaqglpfcrrrmkrkldhgsevrsfslgkkpckvseytsttglvpcsatptt 60

Qy     91 FGDLRAANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHM 150
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 fgdlraangqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhm 120

Qy    151 MQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEE 210
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 mqviepqfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckee 180

Qy    211 GIDEPLHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQ 270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gideplhikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclq 240

Qy    271 FCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNA 330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 fcgnrivsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwna 300

Qy    331 ETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQ 390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 etgecihtlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvq 360

Qy    391 YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDV 450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ydgrrvvsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdv 420

Qy    451 ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVT 510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 etgncihtltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavt 480

Qy    511 CLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGS 570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 clqfnknfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgs 540

Qy    571 RNGTEETKLLVLDFDVDMK 589
          |||||||||||||||||||
Db    541 rngteetkllvldfdvdmk 559


RESULT   4
AAB59199
ID   AAB59199 standard; protein; 559 AA.
XX
AC   AAB59199;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human mammary sel-10-2 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 83-86; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   559 AA;


  Query Match             94.9%;  Score 559;  DB 22;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  559;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     31 MKIFSISIIAQGLPFCRRRMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTT 90
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 mkifsisiiaqglpfcrrrmkrkldhgsevrsfslgkkpckvseytsttglvpcsatptt 60

Qy         91 FGDLRAANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHM 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 fgdlraangqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhm 120

Qy        151 MQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEE 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 mqviepqfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckee 180

Qy        211 GIDEPLHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQ 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 gideplhikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclq 240

Qy        271 FCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNA 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 fcgnrivsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwna 300

Qy        331 ETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQ 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 etgecihtlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvq 360

Qy        391 YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDV 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ydgrrvvsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdv 420

Qy        451 ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVT 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 etgncihtltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavt 480

Qy        511 CLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGS 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 clqfnknfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgs 540

Qy        571 RNGTEETKLLVLDFDVDMK 589
              |||||||||||||||||||
Db        541 rngteetkllvldfdvdmk 559


RESULT    5
AAB01204
ID   AAB01204 standard; Protein; 589 AA.
XX
AC   AAB01204;
XX
DT   03-NOV-2000  (first entry)
XX
DE   Human GTPase associated protein-29.
XX
KW   Guanine nucleotide binding protein; GTP-binding protein; G-protein;
KW   GTPase; GTPase associated protein; GTPAP; cell proliferation;
KW   autoimmune; inflammatory; immune system disorder; cancer; AIDS;
KW   acquired immune deficiency syndrome; asthma; atherosclerosis;
KW   arthritis; systemic lupus erythematosus; psoriasis; human.
XX
OS   Homo sapiens.
XX
PN   WO200031263-A2.
XX
PD   02-JUN-2000.
XX
PF   23-NOV-1999;   99WO-US28013.
XX
PR   23-NOV-1998;   98US-0109592.
PR   04-FEB-1999;   99US-0118610.
PR   06-APR-1999;   99US-0127990.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL,  Tang YT,  Bandman O,  Lal P,  Yue H,  Lu DAM,  Baughn MR;
PI   Yang J,  Azimzai Y;
XX
DR   WPI; 2000-400073/34.
DR   N-PSDB; AAA49199.
XX
PT   Human GTPase associated proteins, polynucleotides, and antibodies,
PT   useful for diagnosing, preventing and treating various diseases such as
PT   atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),
PT   asthma, and autoimmune diseases -
XX
PS   Claim 1; Page 118-120; 144pp; English.
XX
CC   Human cDNA libraries from various tissues were screened for GTPase
CC   associated proteins (GTPAP). The present sequence is human
CC   GTPAP-29 protein. This sequence was derived from a cDNA library of
CC   brain tumour tissue. This protein is expressed in nervous,
CC   reproductive and gastrointestinal tissue. The GTPAP proteins may be
CC   used to define agonists and antagonists of GTPAP activity and to
CC   generate antibodies to GTPAP. This means the GTPAP proteins may be
CC   useful for treatment or prevention of diseases associated with GTPAP
CC   such as cell proliferation disorders, autoimmune disorders,
CC   inflammatory disorders, immune system disorders, cancer, AIDS, asthma,
CC   atherosclerosis, arthritis, systemic lupus erythematosus and psoriasis.
XX
SQ   Sequence   589 AA;

  Query Match             94.6%;  Score 557;  DB 21;  Length 589;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  557;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         33 IFSISIIAQGLPFCRRRMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFG 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 ifsisiiaqglpfcrrrmkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfg 92

Qy         93 DLRAANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQ 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 dlraangqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmq 152

Qy        153 VIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGI 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 viepqfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckeegi 212

Qy        213 DEPLHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFC 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 deplhikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclqfc 272

Qy        273 GNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAET 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 gnrivsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaet 332

Qy        333 GECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYD 392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 gecihtlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqyd 392

Qy        393 GRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVET 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 grrvvsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvet 452

Qy        453 GNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCL 512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 gncihtltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtcl 512

Qy        513 QFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRN 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 qfnknfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrn 572

Qy        573 GTEETKLLVLDFDVDMK 589
              |||||||||||||||||
Db        573 gteetkllvldfdvdmk 589


RESULT    6
AAY22465
ID   AAY22465 standard; Protein; 540 AA.
```

```
XX
AC   AAY22465;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human hippocampal sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99701.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 57-59; 91pp; English.
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human hippocampal sel-10 protein. The
CC   polypeptides can be used to alter presenilin function. Compounds which
CC   inhibit either the expression or the activity of the human sel-10
CC   polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)
CC   or PS-2, and therefore may be useful for the prevention or treatment of
CC   Alzheimer's disease.
XX
SQ   Sequence   540 AA;

  Query Match             91.7%;  Score 540;  DB 20;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 60

Qy     110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 120

Qy     170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 180

Qy     230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 240

Qy     290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 300

Qy     350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 360

Qy     410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 420

Qy     470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 480

Qy     530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 540

RESULT   7
AAY22468
ID   AAY22468 standard; Protein; 540 AA.
XX
AC   AAY22468;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human mammary sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;
KW   mammary gland; therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99702.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 66-69; 91pp; English.
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human mammary sel-10 protein. The polypeptides
CC   can be used to alter presenilin function. Compounds which inhibit either
CC   the expression or the activity of the human sel-10 polypeptides may
CC   reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and
CC   therefore may be useful for the prevention or treatment of Alzheimer's
CC   disease.
XX
SQ   Sequence   540 AA;

  Query Match             91.7%;  Score 540;  DB 20;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 60

Qy     110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 120

Qy     170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 180

Qy     230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    181 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 240

Qy    290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 300

Qy    350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 360

Qy    410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 420

Qy    470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 480

Qy    530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 540


RESULT   8
AAB59197
ID   AAB59197 standard; protein; 540 AA.
XX
AC   AAB59197;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human hippocampal sel-10-5 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 75-78; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   540 AA;

  Query Match             91.7%;  Score 540;  DB 22;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 60

Qy    110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 120

Qy    170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 180

Qy    230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 240

Qy    290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 300

Qy    350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 360

Qy    410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 420

Qy    470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 480

Qy    530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 540


RESULT   9
AAB59200
ID   AAB59200 standard; protein; 540 AA.
XX
AC   AAB59200;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human mammary sel-10-3 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 86-89; 116pp; English.
```

```
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   540 AA;

  Query Match             91.7%;  Score 540;  DB 22;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 60

Qy     110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 120

Qy     170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 180

Qy     230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 240

Qy     290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 300

Qy     350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 360

Qy     410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 420

Qy     470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 480

Qy     530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 540

RESULT  10
AAY22464
ID   AAY22464 standard; Protein; 545 AA.
XX
AC   AAY22464;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human hippocampal sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
```

```
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99701.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 53-56; 91pp; English.
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human hippocampal sel-10 protein. The
CC   polypeptides can be used to alter presenilin function. Compounds which
CC   inhibit either the expression or the activity of the human sel-10
CC   polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)
CC   or PS-2, and therefore may be useful for the prevention or treatment of
CC   Alzheimer's disease.
XX
SQ   Sequence   545 AA;

  Query Match             91.7%;  Score 540;  DB 20;  Length 545;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 65

Qy     110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 125

Qy     170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 185

Qy     230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 245

Qy     290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     246 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 305

Qy     350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 365

Qy     410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 425

Qy     470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     426 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 485

Qy     530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     486 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 545

RESULT  11
AAB59196
ID   AAB59196 standard; protein; 545 AA.
XX
AC   AAB59196;
```

```
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human hippocampal sel-10-4 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 72-75; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   545 AA;

  Query Match             91.7%;  Score 540;  DB 22;  Length 545;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 65

Qy        110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 125

Qy        170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 185

Qy        230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 245

Qy        290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 305

Qy        350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 365

Qy        410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 425

Qy        470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 485

Qy        530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 545


RESULT 12
AAY22463
ID   AAY22463 standard; Protein; 553 AA.
XX
AC   AAY22463;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human hippocampal sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99701.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 50-53; 91pp; English.
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human hippocampal sel-10 protein. The
CC   polypeptides can be used to alter presenilin function. Compounds which
CC   inhibit either the expression or the activity of the human sel-10
CC   polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)
CC   or PS-2, and therefore may be useful for the prevention or treatment of
CC   Alzheimer's disease.
XX
SQ   Sequence   553 AA;

  Query Match             91.7%;  Score 540;  DB 20;  Length 553;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 73

Qy        110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 133

Qy        170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 193

Qy        230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 253

Qy        290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 313

Qy        350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 373

Qy        410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 433

Qy        470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 493

Qy        530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 553


RESULT 13
AAB59195
ID   AAB59195 standard; protein; 553 AA.
XX
AC   AAB59195;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human hippocampal sel-10-3 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 68-71; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   553 AA;

  Query Match             91.7%;  Score 540;  DB 22;  Length 553;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 73

Qy        110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 133

Qy        170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 193

Qy        230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 253

Qy        290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 313

Qy        350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 373

Qy        410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 433

Qy        470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 493

Qy        530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 553


RESULT 14
AAY22462
ID   AAY22462 standard; Protein; 592 AA.
XX
AC   AAY22462;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human hippocampal sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99701.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 47-50; 91pp; English.
```

```
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human hippocampal sel-10 protein. The
CC   polypeptides can be used to alter presenilin function. Compounds which
CC   inhibit either the expression or the activity of the human sel-10
CC   polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)
CC   or PS-2, and therefore may be useful for the prevention or treatment of
CC   Alzheimer's disease.
XX
SQ   Sequence   592 AA;

  Query Match             91.7%;  Score 540;  DB 20;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 112

Qy        110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 172

Qy        170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 232

Qy        230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 292

Qy        290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 352

Qy        350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 412

Qy        410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 472

Qy        470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 532

Qy        530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 592

RESULT 15
AAB59194
ID   AAB59194 standard; protein; 592 AA.
XX
AC   AAB59194;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human hippocampal sel-10-2 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 64-68; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   592 AA;

  Query Match             91.7%;  Score 540;  DB 22;  Length 592;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 mkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrits 112

Qy        110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 vqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpke 172

Qy        170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 lalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfi 232

Qy        230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvws 292

Qy        290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 avtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcm 352

Qy        350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 hlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwd 412

Qy        410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 petetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 472

Qy        470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 elkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvk 532

Qy        530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 lwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 592


Search completed: September 25, 2001, 14:45:34
Job time: 209 sec
```

THIS PAGE BLANK (USPTO)

```
                      GenCore version 4.5
            Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         September 25, 2001, 14:43:40 ; Search time 13.51 Seconds
                                               (without alignments)
                                               897.684 Million cell updates/sec

Title:          US-09-328-877A-8
Perfect score:  589
Sequence:       1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:  OLIGO
                Gapop 60.0 , Gapext 60.0

Searched:       197339 seqs, 20590346 residues

Word size :     0

Total number of hits satisfying chosen parameters:      197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :      Issued_Patents_AA:*
              1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
              2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
              3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
              4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
              5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
              6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES

                    %
Result            Query
   No.    Score   Match Length DB  ID                         Description
------------------------------------------------------------------------------
     1      12     2.0    209  3  US-08-899-578-6              Sequence 6, Appli
     2      12     2.0    587  3  US-08-899-578-2              Sequence 2, Appli
     3       8     1.4     13  2  US-08-671-094B-3             Sequence 3, Appli
     4       8     1.4     13  2  US-08-671-094B-5             Sequence 5, Appli
     5       8     1.4     14  2  US-08-671-094B-2             Sequence 2, Appli
     6       8     1.4     15  1  US-08-311-307B-6             Sequence 6, Appli
     7       8     1.4     16  1  US-08-311-307B-3             Sequence 3, Appli
     8       8     1.4     16  1  US-08-311-307B-4             Sequence 4, Appli
     9       8     1.4     37  3  US-08-476-376-12             Sequence 12, Appl
    10       8     1.4     73  1  US-07-792-988-3              Sequence 3, Appli
    11       8     1.4     73  1  US-07-792-988-4              Sequence 4, Appli
    12       8     1.4     73  1  US-07-778-413E-18            Sequence 18, Appl
    13       8     1.4     73  1  US-07-778-413E-19            Sequence 19, Appl
    14       8     1.4     73  1  US-07-778-413E-20            Sequence 20, Appl
    15       8     1.4     73  1  US-08-340-102-18             Sequence 18, Appl
    16       8     1.4     73  1  US-08-340-102-19             Sequence 19, Appl
    17       8     1.4     73  1  US-08-340-102-20             Sequence 20, Appl
    18       8     1.4     73  1  US-08-330-163-2              Sequence 2, Appli
    19       8     1.4     73  1  US-08-330-163-3              Sequence 3, Appli
    20       8     1.4     73  1  US-08-330-163-4              Sequence 4, Appli
    21       8     1.4     73  1  US-08-482-111-2              Sequence 2, Appli
    22       8     1.4     73  1  US-08-482-111-3              Sequence 3, Appli
    23       8     1.4     73  1  US-08-482-111-4              Sequence 4, Appli
    24       8     1.4     73  2  US-08-436-420-26             Sequence 26, Appl
    25       8     1.4     73  2  US-08-436-420-27             Sequence 27, Appl
    26       8     1.4     73  2  US-08-436-420-28             Sequence 28, Appl
    27       8     1.4     73  3  US-08-846-966-2              Sequence 2, Appli
    28       8     1.4     73  3  US-08-846-966-3              Sequence 3, Appli
    29       8     1.4     73  3  US-08-846-966-4              Sequence 4, Appli
    30       8     1.4     73  3  US-08-557-142-2              Sequence 2, Appli
    31       8     1.4     73  3  US-08-557-142-3              Sequence 3, Appli
    32       8     1.4     73  3  US-08-557-142-4              Sequence 4, Appli
    33       8     1.4     73  5  PCT-US94-06264-2             Sequence 2, Appli
    34       8     1.4     73  5  PCT-US94-06264-3             Sequence 3, Appli
    35       8     1.4     73  5  PCT-US94-06264-4             Sequence 4, Appli
    36       8     1.4    106  1  US-08-352-324A-5             Sequence 5, Appli
    37       8     1.4    106  2  US-08-862-607-5              Sequence 5, Appli
    38       8     1.4    106  2  US-08-468-819-7              Sequence 7, Appli
    39       8     1.4    106  3  US-09-203-235-5              Sequence 5, Appli
    40       8     1.4    106  5  PCT-US95-16144-5             Sequence 5, Appli
    41       8     1.4    107  1  US-08-352-324A-4             Sequence 4, Appli
    42       8     1.4    107  1  US-08-352-324A-7             Sequence 7, Appli
    43       8     1.4    107  2  US-08-862-607-4              Sequence 4, Appli
    44       8     1.4    107  2  US-08-862-607-7              Sequence 7, Appli
    45       8     1.4    107  2  US-08-468-819-5              Sequence 5, Appli


                                ALIGNMENTS


RESULT   1
US-08-899-578-6
; Sequence 6, Application US/08899578
; Patent No. 6087153
;  GENERAL INFORMATION:
;    APPLICANT:  Greenwald, Iva
;    APPLICANT:  Hubbard, E. Jane
;    TITLE OF INVENTION:  SEL-10 AND USES THEREOF
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/899,578
;      FILING DATE:  24-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/53200/JPW/AKC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 278-0525
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  209 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-899-578-6


  Query Match             2.0%;  Score 12;  DB 3;  Length 209;
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     438 SGSLDTSIRVWD 449
           ||||||||||||
Db     139 SGSLDTSIRVWD 150
```

```
RESULT   2
US-08-899-578-2
; Sequence 2, Application US/08899578
; Patent No. 6087153
;  GENERAL INFORMATION:
;    APPLICANT:  Greenwald, Iva
;    APPLICANT:  Hubbard, E. Jane
;    TITLE OF INVENTION:  SEL-10 AND USES THEREOF
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/899,578
;      FILING DATE:  24-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/53200/JPW/AKC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 278-0525
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 587 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-899-578-2


  Query Match             2.0%;  Score 12;  DB 3;  Length 587;
  Best Local Similarity   100.0%;  Pred. No. 0.004;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    438 SGSLDTSIRVWD 449
          ||||||||||||
Db    435 SGSLDTSIRVWD 446


RESULT   3
US-08-671-094B-3
; Sequence 3, Application US/08671094B
; Patent No. 5912232
;  GENERAL INFORMATION:
;    APPLICANT:  Talmadge, James E.
;    TITLE OF INVENTION:  Anti-inflammatory Polypeptide
;    TITLE OF INVENTION:  Antagonists of Human Interleukin-8
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dann, Dorfman, Herrell and Skillman
;      STREET:  1601 Market Street
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-2307
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/671,094B
;      FILING DATE:  28-JUN-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hagan, Patrick J.
;      REGISTRATION NUMBER:  27,643
;      REFERENCE/DOCKET NUMBER:  63086FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215)563-4100
;      TELEFAX:  (215)563-4044
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4..6
;      OTHER INFORMATION:  /note= "The cysteine residues
;      OTHER INFORMATION:  may be substituted with aminobutryic acid, homocysteine, o
;      OTHER INFORMATION:  diaminosuberic acid."
US-08-671-094B-3


  Query Match             1.4%;  Score 8;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.76;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db      3 QCLQTLQG 10


RESULT   4
US-08-671-094B-5
; Sequence 5, Application US/08671094B
; Patent No. 5912232
;  GENERAL INFORMATION:
;    APPLICANT:  Talmadge, James E.
;    TITLE OF INVENTION:  Anti-inflammatory Polypeptide
;    TITLE OF INVENTION:  Antagonists of Human Interleukin-8
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dann, Dorfman, Herrell and Skillman
;      STREET:  1601 Market Street
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-2307
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/671,094B
;      FILING DATE:  28-JUN-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hagan, Patrick J.
;      REGISTRATION NUMBER:  27,643
;      REFERENCE/DOCKET NUMBER:  63086FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215)563-4100
;      TELEFAX:  (215)563-4044
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4..6
;      OTHER INFORMATION:  /note= "The cysteine residues
;      OTHER INFORMATION:  may be substituted with aminobutyric acid, homocysteine, or
;      OTHER INFORMATION:  diaminosuberic acid."
US-08-671-094B-5


  Query Match              1.4%;  Score 8;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.76;
  Matches     8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     494 QCLQTLQG 501
           ||||||||
Db       3 QCLQTLQG 10


RESULT   5
US-08-671-094B-2
; Sequence 2, Application US/08671094B
; Patent No. 5912232
;  GENERAL INFORMATION:
;    APPLICANT:  Talmadge, James E.
;    TITLE OF INVENTION:  Anti-inflammatory Polypeptide
;    TITLE OF INVENTION:  Antagonists of Human Interleukin-8
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dann, Dorfman, Herrell and Skillman
;      STREET:  1601 Market Street
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-2307
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/671,094B
;      FILING DATE:  28-JUN-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hagan, Patrick J.
;      REGISTRATION NUMBER:  27,643
;      REFERENCE/DOCKET NUMBER:  63086FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215)563-4100
;      TELEFAX:  (215)563-4044
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4..6
;      OTHER INFORMATION:  /note= "The cysteine residues
;      OTHER INFORMATION:  may be substituted with aminobutyric acid, homocysteine, or
;      OTHER INFORMATION:  diaminosuberic acid."
US-08-671-094B-2
```

```
  Query Match              1.4%;  Score 8;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.81;
  Matches     8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     494 QCLQTLQG 501
           ||||||||
Db       3 QCLQTLQG 10


RESULT   6
US-08-311-307B-6
; Sequence 6, Application US/08311307B
; Patent No. 5627156
;  GENERAL INFORMATION:
;    APPLICANT:  Talmadge, James E.
;    TITLE OF INVENTION:  Polypeptide Agonist Derived From Human
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dann, Dorfman, Herrell and Skillman
;      STREET:  1601 Market Street
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-2307
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,307B
;      FILING DATE:  23-SEP-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hagan, Patrick J.
;      REGISTRATION NUMBER:  27,643
;      REFERENCE/DOCKET NUMBER:  63085
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215)563-4100
;      TELEFAX:  (215)563-4044
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4..6
;      OTHER INFORMATION:  /note= "The cysteine residues may
;      OTHER INFORMATION:  be substituted with aminobutyric acid, homocysteine, or
;      OTHER INFORMATION:  diaminosuberic acid."
US-08-311-307B-6


  Query Match              1.4%;  Score 8;  DB 1;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.86;
  Matches     8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     494 QCLQTLQG 501
           ||||||||
Db       5 QCLQTLQG 12


RESULT   7
US-08-311-307B-3
; Sequence 3, Application US/08311307B
; Patent No. 5627156
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Talmadge, James E.
;    TITLE OF INVENTION:  Polypeptide Agonist Derived From Human
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dann, Dorfman, Herrell and Skillman
;      STREET:  1601 Market Street
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-2307
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,307B
;      FILING DATE:  23-SEP-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hagan, Patrick J.
;      REGISTRATION NUMBER:  27,643
;      REFERENCE/DOCKET NUMBER:  63085
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215)563-4100
;      TELEFAX:  (215)563-4044
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4..6
;      OTHER INFORMATION:  /note= "The cysteine residues may
;      OTHER INFORMATION:  be substituted with aminobutryic acid, homocysteine, or
;      OTHER INFORMATION:  diaminosuberic acid."
US-08-311-307B-3


  Query Match             1.4%;  Score 8;  DB 1;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.91;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db      5 QCLQTLQG 12


RESULT   8
US-08-311-307B-4
; Sequence 4, Application US/08311307B
; Patent No. 5627156
;  GENERAL INFORMATION:
;    APPLICANT:  Talmadge, James E.
;    TITLE OF INVENTION:  Polypeptide Agonist Derived From Human
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dann, Dorfman, Herrell and Skillman
;      STREET:  1601 Market Street
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-2307
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,307B
;      FILING DATE:  23-SEP-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hagan, Patrick J.
;      REGISTRATION NUMBER:  27,643
;      REFERENCE/DOCKET NUMBER:  63085
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215)563-4100
;      TELEFAX:  (215)563-4044
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4..6
;      OTHER INFORMATION:  /note= "The cysteine residues may
;      OTHER INFORMATION:  be substituted with aminobutyric acid, homocysteine or
;      OTHER INFORMATION:  diaminosuberic acid."
US-08-311-307B-4


  Query Match             1.4%;  Score 8;  DB 1;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.91;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db      5 QCLQTLQG 12


RESULT   9
US-08-476-376-12
; Sequence 12, Application US/08476376
; Patent No. 6103234
;  GENERAL INFORMATION:
;    APPLICANT:  WOLPE, STEPHEN D.
;    APPLICANT:  CERAMI, ANTHONY
;    APPLICANT:  SHERRY, BARBARA
;    APPLICANT:  TEKAMP-OLSON, PATRICIA A.
;    TITLE OF INVENTION:  MACROPHAGE-DERIVED INFLAMMATORY MEDIATOR
;    TITLE OF INVENTION:  (MIP-2)
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DAVID A. JACKSON, ESQ.
;      STREET:  411 HACKENSACK AVE, CONTINENTAL PLAZA, 4TH
;      STREET:  FLOOR
;      CITY:  HACKENSACK
;      STATE:  NEW JERSEY
;      COUNTRY:  U.S.A.
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,376
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/285,498
;      FILING DATE:  03-AUG-1994
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 08/105,105
;      FILING DATE:  10-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/914,045
;      FILING DATE:  13-JUL-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/399,971
;      FILING DATE:  01-SEP-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/240,078
;      FILING DATE:  02-SEP-1988
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/104,827
;      FILING DATE:  02-OCT-1987
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 06/766,852
;      FILING DATE:  16-AUG-1985
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 06/414,098
;      FILING DATE:  07-SEP-1982
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 06/351,290
;      FILING DATE:  22-FEB-1982
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 06/299,932
;      FILING DATE:  08-SEP-1981
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JACKSON ESQ., DAVID A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  13046 CIP A2D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  37 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-476-376-12

  Query Match              1.4%;  Score 8;  DB 3;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     15 QCLQTLQG 22


RESULT  10
US-07-792-988-3
; Sequence 3, Application US/07792988
; Patent No. 5306709
;  GENERAL INFORMATION:
;    APPLICANT:  Gewirtz, Alan M.
;    TITLE OF INVENTION:  Suppression of megakaryo-
;    TITLE OF INVENTION:  cytopoiesis by macrophage inflammatory proteins
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  University of Pennsylvania
;      STREET:  Suite 419
;      STREET:  133 South 36th Street
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19104-3246
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/792,988
;      FILING DATE:  19911115
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-159
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5306709e
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  73 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single stranded
;      TOPOLOGY:  linear
US-07-792-988-3

  Query Match              1.4%;  Score 8;  DB 1;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     10 QCLQTLQG 17


RESULT  11
US-07-792-988-4
; Sequence 4, Application US/07792988
; Patent No. 5306709
;  GENERAL INFORMATION:
;    APPLICANT:  Gewirtz, Alan M.
;    TITLE OF INVENTION:  Suppression of megakaryo-
;    TITLE OF INVENTION:  cytopoiesis by macrophage inflammatory proteins
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  University of Pennsylvania
;      STREET:  Suite 419
;      STREET:  133 South 36th Street
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19104-3246
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/792,988
;      FILING DATE:  19911115
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-159
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5306709e
```

```
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  73 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single stranded
;      TOPOLOGY:  linear
US-07-792-988-4


  Query Match              1.4%;  Score 8;  DB 1;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     10 QCLQTLQG 17


RESULT  12
US-07-778-413E-18
; Sequence 18, Application US/07778413E
; Patent No. 5401651
;  GENERAL INFORMATION:
;    APPLICANT:  Walz, Alfred
;    TITLE OF INVENTION:  No. 5401651el Neutrophil
;    TITLE OF INVENTION:  Activating Factors
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  Amgen Center
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in., DS, 2.0 MB
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Macintosh OS 7.0
;      SOFTWARE:  Microsoft Word Version 5.1a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/778,413E
;      FILING DATE:  16-OCT-1991
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cook, Robert R.
;      REGISTRATION NUMBER:  31602
;      REFERENCE/DOCKET NUMBER:  A-204
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (805) 499-5725 EXTENSION 4955
;      TELEFAX:  (805) 499-8011
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  73 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-778-413E-18


  Query Match              1.4%;  Score 8;  DB 1;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     10 QCLQTLQG 17


RESULT  13
US-07-778-413E-19
; Sequence 19, Application US/07778413E
; Patent No. 5401651
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Walz, Alfred
;    TITLE OF INVENTION:  No. 5401651el Neutrophil
;    TITLE OF INVENTION:  Activating Factors
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  Amgen Center
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in., DS, 2.0 MB
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Macintosh OS 7.0
;      SOFTWARE:  Microsoft Word Version 5.1a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/778,413E
;      FILING DATE:  16-OCT-1991
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cook, Robert R.
;      REGISTRATION NUMBER:  31602
;      REFERENCE/DOCKET NUMBER:  A-204
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (805) 499-5725 EXTENSION 4955
;      TELEFAX:  (805) 499-8011
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  73 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-778-413E-19


  Query Match              1.4%;  Score 8;  DB 1;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     10 QCLQTLQG 17


RESULT  14
US-07-778-413E-20
; Sequence 20, Application US/07778413E
; Patent No. 5401651
;  GENERAL INFORMATION:
;    APPLICANT:  Walz, Alfred
;    TITLE OF INVENTION:  No. 5401651el Neutrophil
;    TITLE OF INVENTION:  Activating Factors
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  Amgen Center
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in., DS, 2.0 MB
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Macintosh OS 7.0
;      SOFTWARE:  Microsoft Word Version 5.1a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/778,413E
;      FILING DATE:  16-OCT-1991
;      CLASSIFICATION:  536
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cook, Robert R.
;      REGISTRATION NUMBER:  31602
;      REFERENCE/DOCKET NUMBER:  A-204
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (805) 499-5725 EXTENSION 4955
;      TELEFAX:  (805) 499-8011
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  73 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-778-413E-20


  Query Match             1.4%;  Score 8;  DB 1;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     10 QCLQTLQG 17


RESULT  15
US-08-340-102-18
; Sequence 18, Application US/08340102
; Patent No. 5591718
;  GENERAL INFORMATION:
;    APPLICANT:  Walz, Alfred
;    TITLE OF INVENTION:  No. 5591718el Neutrophil
;    TITLE OF INVENTION:  Activating Factors
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  Amgen Center
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in., DS, 1.4 MB
;      COMPUTER:  MS-DOS
;      OPERATING SYSTEM:  MS-DOS 6.22
;      SOFTWARE:  Microsoft Word Version 5.1a for
;      SOFTWARE:  Macintosh
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,102
;      FILING DATE:  15-NOV-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cook, Robert R.
;      REGISTRATION NUMBER:  31602
;      REFERENCE/DOCKET NUMBER:  A-204A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (805) 499-5725 EXTENSION 4955
;      TELEFAX:  (805) 499-8011
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  73 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-340-102-18


  Query Match             1.4%;  Score 8;  DB 1;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 3.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     10 QCLQTLQG 17


Search completed: September 25, 2001, 14:45:58
Job time: 138 sec
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         September 25, 2001, 14:44:10 ; Search time 17.49 Seconds
                                                (without alignments)
                                                2565.284 Million cell updates/sec

Title:          US-09-328-877A-8
Perfect score:  589
Sequence:       1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:  OLIGO
                Gapop 60.0 , Gapext 60.0

Searched:       219241 seqs, 76174552 residues

Word size :     0

Total number of hits satisfying chosen parameters:      219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :      PIR_68:*
               1:  pir1:*
               2:  pir2:*
               3:  pir3:*
               4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                     %
Result          Query
  No.   Score   Match Length DB  ID                   Description
-------------------------------------------------------------------------------
    1     12    2.0    579  2  T22703               hypothetical prote
    2      9    1.5    524  2  T46310               hypothetical prote
    3      8    1.4     33  2  S50035               cytokine boGRO - b
    4      8    1.4    107  2  B38290               GRO-gamma precurso
    5      8    1.4    107  2  A28414               melanoma growth-st
    6      8    1.4    107  2  JH0281               macrophage inflamm
    7      8    1.4    272  2  H70335               hypothetical prote
    8      8    1.4    330  2  S27707               daunorubicin resis
    9      8    1.4    336  2  T27762               hypothetical prote
   10      8    1.4    344  2  I45774               odorant receptor 1
   11      8    1.4    386  2  T26805               hypothetical prote
   12      8    1.4    515  2  A43289               CDC4L protein - hu
   13      8    1.4    701  2  T16607               hypothetical prote
   14      8    1.4    821  2  A86243               hypothetical prote
   15      7    1.2     40  2  S56768               capsid protein - L
   16      7    1.2    107  2  JQ2035               hypothetical 11.7K
   17      7    1.2    123  2  S29121               GTP-binding regula
   18      7    1.2    135  2  T03288               ferredoxin [2Fe-2S
   19      7    1.2    149  2  S35208               serine proteinase
   20      7    1.2    149  2  T23179               hypothetical prote
   21      7    1.2    164  2  T33915               hypothetical prote
   22      7    1.2    187  2  A72737               hypothetical prote
   23      7    1.2    190  2  T35570               hypothetical prote
   24      7    1.2    196  2  T26943               hypothetical prote
   25      7    1.2    206  2  H81932               probable dTMP kina
   26      7    1.2    206  2  B81173               thymidylate kinase
   27      7    1.2    209  2  G71430               hypothetical prote
   28      7    1.2    221  2  T10221               hypothetical prote
   29      7    1.2    227  2  T21783               hypothetical prote
   30      7    1.2    232  2  T21359               hypothetical prote
   31      7    1.2    249  2  A96632               hypothetical prote
   32      7    1.2    251  2  B70450               beta 1,4 glucosylt
   33      7    1.2    253  2  T16918               hypothetical prote
   34      7    1.2    255  2  S55641               uracil DNA glycosy
   35      7    1.2    268  2  S76063               hypothetical prote
   36      7    1.2    273  2  H86241               protein T16B5.4 [i
   37      7    1.2    273  2  S61637               probable membrane
   38      7    1.2    289  2  H82364               O-methyltransferas
   39      7    1.2    298  2  C83216               probable epoxide h
   40      7    1.2    304  2  T40316               conserved kin7-lik
   41      7    1.2    310  2  T43158               probable GTP-bindi
   42      7    1.2    314  2  T43299               probable GTP-bindi
   43      7    1.2    319  2  A70340               conserved hypothet
   44      7    1.2    319  2  A72637               hypothetical prote
   45      7    1.2    326  1  RGBOB2               GTP-binding regula

                              ALIGNMENTS

RESULT   1
T22703
hypothetical protein F55B12.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22703
R;Sims, M.
submitted to the EMBL Data Library, September 1996
A;Reference number: Z19602
A;Accession: T22703
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-579 <WIL>
A;Cross-references: EMBL:Z79757; PIDN:CAB02129.1; GSPDB:GN00023; CESP:F55B12.3
A;Experimental source: clone F55B12
C;Genetics:
A;Gene: CESP:F55B12.3
A;Map position: 5
A;Introns: 22/3; 45/2; 77/3; 122/3; 171/2; 211/2; 342/2; 417/3; 478/3; 526/1

  Query Match             2.0%;  Score 12;  DB 2;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 0.00071;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     438 SGSLDTSIRVWD 449
           ||||||||||||
Db     427 SGSLDTSIRVWD 438


RESULT   2
T46310
hypothetical protein DKFZp434G0511.1 - human
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46310
R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23035
A;Accession: T46310
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-524 <AAA>
A;Cross-references: EMBL:AL137748
A;Experimental source: adult testis; clone DKFZp434G0511
C;Genetics:
A;Note: DKFZp434G0511.1


  Query Match             1.5%;  Score 9;  DB 2;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 0.91;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy    357 VSGSRDATL 365
          |||||||||
Db    318 VSGSRDATL 326


RESULT   3
S50035
cytokine boGRO - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 13-Jan-1995 #sequence_revision 12-May-1995 #text_change 02-Jun-1995
C;Accession: S50035
R;Rogivue, C.; Car, B.D.; Allmann-Iselin, I.; Zwahlen, R.D.; Walz, A.
submitted to the Protein Sequence Database, December 1994
A;Description: Bovine GRO (boGRO), a new monocyte-macrophage-derived cytokine of the int
A;Reference number: S50035
A;Accession: S50035
A;Molecule type: protein
A;Residues: 1-33 <ROG>
C;Superfamily: beta-thromboglobulin


  Query Match             1.4%;  Score 8;  DB 2;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 0.88;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     10 QCLQTLQG 17


RESULT   4
B38290
GRO-gamma precursor - human
N;Alternate names: growth-regulated protein gamma; macrophage inflammatory protein 2 bet
C;Species: Homo sapiens (man)
C;Date: 31-May-1991 #sequence_revision 27-Oct-1995 #text_change 20-Aug-1999
C;Accession: JH0282; B38290; C46519
R;Tekamp-Olson, P.; Gallegos, C.; Bauer, D.; McClain, J.; Sherry, B.; Fabre, M.; van Dev
J. Exp. Med. 172, 911-919, 1990
A;Title: Cloning and characterization of cDNAs for murine macrophage inflammatory protei
A;Reference number: JH0200; MUID:90354792
A;Accession: JH0282
A;Molecule type: mRNA
A;Residues: 1-107 <TEK>
A;Cross-references: GB:X53800; NID:g34662; PIDN:CAA37809.1; PID:g34663
R;Haskill, S.; Peace, A.; Morris, J.; Sporn, S.A.; Anisowicz, A.; Lee, S.W.; Smith, T.;
Proc. Natl. Acad. Sci. U.S.A. 87, 7732-7736, 1990
A;Title: Identification of three related human GRO genes encoding cytokine functions.
A;Reference number: A38290; MUID:91017578
A;Accession: B38290
A;Molecule type: mRNA
A;Residues: 1-26,'G',29-107 <HAS>
A;Cross-references: GB:M36821; NID:g183632; PIDN:AAA63184.1; PID:g183633
R;Proost, P.; De Wolf-Peeters, C.; Conings, R.; Opdenakker, G.; Billiau, A.; Van Damme,
J. Immunol. 150, 1000-1010, 1993
A;Title: Identification of a novel granulocyte chemotactic protein (GCP-2) from human tu
A;Reference number: A46519; MUID:93139489
A;Accession: C46519
A;Molecule type: protein
A;Residues: 35-52 <PRO>
A;Experimental source: MG-63 osteosarcoma cells
C;Genetics:
A;Map position: 4q21
C;Superfamily: beta-thromboglobulin
F;1-34/Domain: signal sequence #status predicted <SIG>
F;35-107/Product: GRO-gamma #status experimental <MAT>


  Query Match             1.4%;  Score 8;  DB 2;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy    494 QCLQTLQG 501
          ||||||||
Db     44 QCLQTLQG 51


RESULT   5
A28414
melanoma growth-stimulatory activity precursor - human
N;Alternate names: fibroblast-derived neutrophil-activating protein gamma; GRO-alpha;
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 20-Aug-1999
C;Accession: S13669; A28414; S00983; B60401; S03976; A47626; B46519
R;Baker, N.E.; Kucera, G.; Richmond, A.
Nucleic Acids Res. 18, 6453, 1990
A;Title: Nucleotide sequence of the human melanoma growth stimulatory activity (MGSA)
A;Reference number: S13669; MUID:91057157
A;Accession: S13669
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-107 <BAK>
A;Cross-references: EMBL:X54489; NID:g34625; PIDN:CAA38361.1; PID:g34626
R;Anisowicz, A.; Bardwell, L.; Sager, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 7188-7192, 1987
A;Title: Constitutive overexpression of a growth-regulated gene in transformed Chines
A;Reference number: A94184; MUID:88041072
A;Accession: A28414
A;Molecule type: mRNA
A;Residues: 1-107 <ANI>
A;Cross-references: GB:J03561; NID:g183622; PIDN:AAA35933.1; PID:g306806
R;Richmond, A.; Balentien, E.; Thomas, H.G.; Flaggs, G.; Barton, D.E.; Spiess, J.; Bo
EMBO J. 7, 2025-2033, 1988
A;Title: Molecular characterization and chromosomal mapping of melanoma growth stimul
A;Reference number: S00983; MUID:88328991
A;Accession: S00983
A;Molecule type: mRNA
A;Residues: 1-107 <RIC>
A;Cross-references: EMBL:X12510; NID:g34621; PIDN:CAA31027.1; PID:g34622
R;Schroeder, J.M.; Sticherling, M.; Henneicke, H.H.; Preissner, W.C.; Christophers, E
J. Immunol. 144, 2223-2232, 1990
A;Title: IL-1alpha or tumor necrosis factor-alpha stimulate release of three NAP-1/IL
A;Reference number: A60401; MUID:90187866
A;Accession: B60401
A;Molecule type: protein
A;Residues: 35-42,'X',44,'X',46-48 <SCH>
A;Experimental source: dermal fibroblasts
R;Golds, E.E.; Mason, P.; Nyirkos, P.
Biochem. J. 259, 585-588, 1989
A;Title: Inflammatory cytokines induce synthesis and secretion of gro protein and a n
A;Reference number: S03975; MUID:89246368
A;Accession: S03976
A;Molecule type: protein
A;Residues: 35-41,'X',43-49,'X',51-52,'XX',55-57 <GOL>
R;Schroeder, J.M.; Persoon, N.L.M.; Christophers, E.
J. Exp. Med. 171, 1091-1100, 1990
A;Title: Lipopolysaccharide-stimulated human monocytes secrete, apart from neutrophil
ntity with melanoma growth stimulatory activity.
A;Reference number: A47626; MUID:90217938
A;Accession: A47626
A;Molecule type: protein
A;Residues: 35-63,'X',65 <SC2>
A;Experimental source: LPS-stimulated monocytes
R;Proost, P.; De Wolf-Peeters, C.; Conings, R.; Opdenakker, G.; Billiau, A.; Van Damm
J. Immunol. 150, 1000-1010, 1993
A;Title: Identification of a novel granulocyte chemotactic protein (GCP-2) from human
A;Reference number: A46519; MUID:93139489
A;Accession: B46519
A;Molecule type: protein
A;Residues: 35-62 <PRO>
A;Experimental source: MG-63 osteosarcoma cells
C;Genetics:
A;Gene: GDB:GRO1
A;Cross-references: GDB:120181; OMIM:155730
A;Map position: 4q21-4q21
```

```
C;Superfamily: beta-thromboglobulin
F;1-34/Domain: signal sequence #status predicted <SIG>
F;35-107/Product: melanoma growth-stimulatory activity #status experimental <MAT>


  Query Match             1.4%;  Score 8;  DB 2;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     44 QCLQTLQG 51


RESULT   6
JH0281
macrophage inflammatory protein 2 alpha precursor - human
N;Alternate names: gro-beta; growth regulated protein beta; melanoma growth-stimulatory
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 20-Aug-1999
C;Accession: JH0281; A35931; A38290; A60407
R;Tekamp-Olson, P.; Gallegos, C.; Bauer, D.; McClain, J.; Sherry, B.; Fabre, M.; van Dev
J. Exp. Med. 172, 911-919, 1990
A;Title: Cloning and characterization of cDNAs for murine macrophage inflammatory protei
A;Reference number: JH0200; MUID:90354792
A;Accession: JH0281
A;Molecule type: mRNA
A;Residues: 1-107 <TEK>
A;Cross-references: GB:X53799; NID:g34658; PIDN:CAA37808.1; PID:g34659
R;Iida, N.; Grotendorst, G.R.
Mol. Cell. Biol. 10, 5596-5599, 1990
A;Title: Cloning and sequencing of a new gro transcript from activated human monocytes:
A;Reference number: A35931; MUID:90377259
A;Accession: A35931
A;Molecule type: mRNA
A;Residues: 1-107 <IID>
A;Cross-references: GB:M57731; GB:M36964; NID:g183626; PIDN:AAA63182.1; PID:g183627
R;Haskill, S.; Peace, A.; Morris, J.; Sporn, S.A.; Anisowicz, A.; Lee, S.W.; Smith, T.;
Proc. Natl. Acad. Sci. U.S.A. 87, 7732-7736, 1990
A;Title: Identification of three related human GRO genes encoding cytokine functions.
A;Reference number: A38290; MUID:91017578
A;Accession: A38290
A;Molecule type: mRNA
A;Residues: 1-107 <HAS>
A;Cross-references: GB:M36820; NID:g183628; PIDN:AAA63183.1; PID:g183629
R;Sporn, S.A.; Eierman, D.F.; Johnson, C.E.; Morris, J.; Martin, G.; Ladner, M.; Haskill
J. Immunol. 144, 4434-4441, 1990
A;Title: Monocyte adherence results in selective induction of novel genes sharing homolo
A;Reference number: A60407; MUID:90257367
A;Accession: A60407
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 56-107 <SPO>
C;Superfamily: beta-thromboglobulin
C;Keywords: inflammation
F;1-34/Domain: signal sequence #status predicted <SIG>
F;35-107/Product: macrophage inflammatory protein 2 alpha #status predicted <MAT>


  Query Match             1.4%;  Score 8;  DB 2;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     44 QCLQTLQG 51


RESULT   7
H70335
hypothetical protein aq_397 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C;Accession: H70335
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666
A;Accession: H70335
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-272 <AQF>
A;Cross-references: GB:AE000688; NID:g2983063; PIDN:AAC06680.1; PID:g2983073; GB:AE00
A;Experimental source: strain VF5
C;Genetics:
A;Gene: aq_397
C;Superfamily: Aquifex aeolicus hypothetical protein aq_397


  Query Match             1.4%;  Score 8;  DB 2;  Length 272;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    407 VWDPETET 414
          ||||||||
Db    161 VWDPETET 168


RESULT   8
S27707
daunorubicin resistance protein - Streptomyces peucetius
C;Species: Streptomyces peucetius
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 02-Feb-2001
C;Accession: S27707
R;Guilfolie, P.G.; Hutchinson, C.R.
submitted to the EMBL Data Library, August 1991
A;Description: A bacterial analog of the mdr gene of mammalian tumor cells is present
A;Reference number: S27706
A;Accession: S27707
A;Molecule type: DNA
A;Residues: 1-330 <GUI>
A;Cross-references: EMBL:M73758; NID:g153228; PIDN:AAA74717.1; PID:g153230
C;Genetics:
A;Gene: drrA
A;Start codon: GTG
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog
C;Keywords: antibiotic resistance; ATP; nucleotide binding; P-loop
F;24-215/Domain: ATP-binding cassette homology <ABC>
F;41-48/Region: nucleotide-binding motif A (P-loop)


  Query Match             1.4%;  Score 8;  DB 2;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 6.7;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     59 EVRSFSLG 66
          ||||||||
Db    292 EVRSFSLG 299


RESULT   9
T27762
hypothetical protein ZK177.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27762
R;Anderson, K.
submitted to the EMBL Data Library, July 1995
A;Description: The sequence of C. elegans cosmid ZK177.
A;Reference number: Z20416
A;Accession: T27762
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-336 <AND>
A;Cross-references: EMBL:U21321; PIDN:AAB36970.1; GSPDB:GN00020; CESP:ZK177.6
```

```
A;Experimental source: strain Bristol N2; clone ZK177
C;Genetics:
A;Gene: CESP:ZK177.6
A;Map position: 2
A;Introns: 44/1; 75/3; 103/1; 185/3; 283/1


  Query Match             1.4%;  Score 8;  DB 2;  Length 336;
  Best Local Similarity   100.0%;  Pred. No. 6.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    525 DGTVKLWD 532
          ||||||||
Db    219 DGTVKLWD 226


RESULT  10
I45774
odorant receptor 1 - channel catfish
C;Species: Ictalurus punctatus (channel catfish)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 04-Sep-1998
C;Accession: I45774
R;Ngai, J.; Dowling, M.M.; Buck, L.; Axel, R.; Chess, A.
Cell 72, 657-666, 1993
A;Title: The family of genes encoding odorant receptors in the channel catfish.
A;Reference number: A45774; MUID:93201590
A;Accession: I45774
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-344 <NGA>
A;Experimental source: olfactory epithelium
A;Note: sequence extracted from NCBI backbone (NCBIP:127744)
C;Superfamily: olfactory receptor OR14
C;Keywords: olfaction; transmembrane protein


  Query Match             1.4%;  Score 8;  DB 2;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 7;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     33 IFSISIIA 40
          ||||||||
Db    157 IFSISIIA 164


RESULT  11
T26805
hypothetical protein Y41C4A.11 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26805
R;Steward, C.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z20269
A;Accession: T26805
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-386 <WIL>
A;Cross-references: EMBL:AL032627; PIDN:CAA21549.1; CESP:Y41C4A.11
A;Experimental source: clone Y41C4A
C;Genetics:
A;Gene: CESP:Y41C4A.11
A;Introns: 24/3; 40/3; 100/1; 146/3; 306/3; 350/2


  Query Match             1.4%;  Score 8;  DB 2;  Length 386;
  Best Local Similarity   100.0%;  Pred. No. 7.7;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    275 RIVSGSDD 282
          ||||||||
Db    290 RIVSGSDD 297
```

```
RESULT  12
A43289
CDC4L protein - human (fragment)
N;Alternate names: cell division control protein CDC4 homolog
C;Species: Homo sapiens (man)
C;Date: 27-Apr-1993 #sequence_revision 26-Jul-1996 #text_change 15-Aug-1997
C;Accession: A43289; S30330
R;Feuchter, A.E.; Freeman, J.D.; Mager, D.L.
Genomics 13, 1237-1246, 1992
A;Title: Strategy for detecting cellular transcripts promoted by human endogenous lon
A;Reference number: A43289; MUID:92372019
A;Accession: A43289
A;Molecule type: mRNA
A;Residues: 1-515 <FEU>
A;Cross-references: EMBL:M83822
A;Note: sequence extracted from NCBI backbone (NCBIP:111997)
C;Genetics:
A;Gene: GDB:CDC4L
A;Cross-references: GDB:131404
A;Map position: 19p13.3-19p13.3


  Query Match             1.4%;  Score 8;  DB 2;  Length 515;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    358 SGSRDATL 365
          ||||||||
Db    482 SGSRDATL 489


RESULT  13
T16607
hypothetical protein K10B2.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16607
R;Miller, N.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of C. elegans cosmid K10B2.
A;Reference number: Z18545
A;Accession: T16607
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-701 <MIL>
A;Cross-references: EMBL:U28730; NID:g860694; PID:g860695; PIDN:AAA68258.1; CESP:K10B
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:K10B2.1
A;Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3


  Query Match             1.4%;  Score 8;  DB 2;  Length 701;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    446 RVWDVETG 453
          ||||||||
Db    321 RVWDVETG 328


RESULT  14
A86243
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: A86243
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
```

```
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: A86243
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-821 <STO>
A;Cross-references: GB:AE005172; NID:g1931650; PIDN:AAB65485.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1


  Query Match             1.4%;  Score 8;  DB 2;  Length 821;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     520 ITSSDDGT 527
           ||||||||
Db     426 ITSSDDGT 433


RESULT  15
S56768
capsid protein - Leishmania RNA virus 2 (fragment)
C;Species: Leishmania RNA virus 2
C;Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997
C;Accession: S56768
R;Widmer, G.; Dooley, S.
Nucleic Acids Res. 23, 2300-2304, 1995
A;Title: Phylogenetic analysis of Leishmania RNA virus and Leishmania suggests ancient v
A;Reference number: S56768; MUID:95334386
A;Accession: S56768
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-40 <WID>
A;Cross-references: EMBL:L41164


  Query Match             1.2%;  Score 7;  DB 2;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     393 GRRVVSG 399
           |||||||
Db      11 GRRVVSG 17


Search completed: September 25, 2001, 14:46:28
Job time: 138 sec
```

```
                        GenCore version 4.5
               Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          September 25, 2001, 14:46:01 ; Search time 10.7 Seconds
                                                (without alignments)
                                                1885.652 Million cell updates/sec

Title:           US-09-328-877A-8
Perfect score:   589
Sequence:        1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:   OLIGO
                 Gapop 60.0 , Gapext 60.0

Searched:        93435 seqs, 34255486 residues

Word size :      0

Total number of hits satisfying chosen parameters:      93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :       SwissProt_39:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                      %
Result              Query
  No.      Score    Match  Length DB  ID                           Description
-------------------------------------------------------------------------------
    1        12      2.0     579  1  SE10_CAEEL                  Q93794 caenorhabdi
    2         8      1.4      98  1  GROG_BOVIN                  O46675 bos taurus
    3         8      1.4     104  1  GRO2_RABIT                  P47854 oryctolagus
    4         8      1.4     104  1  GROA_BOVIN                  O46676 bos taurus
    5         8      1.4     104  1  GROB_BOVIN                  O46677 bos taurus
    6         8      1.4     107  1  GRO_HUMAN                   P09341 homo sapien
    7         8      1.4     107  1  MI2A_HUMAN                  P19875 homo sapien
    8         8      1.4     107  1  MI2B_HUMAN                  P19876 homo sapien
    9         8      1.4     330  1  DRRA_STRPE                  P32010 streptomyce
   10         8      1.4     701  1  YSS1_CAEEL                  Q09990 caenorhabdi
   11         8      1.4    1918  1  CC4H_HUMAN                  P50851 homo sapien
   12         7      1.2     135  1  FER5_MAIZE                  P27789 zea mays (m
   13         7      1.2     170  1  DCB1_METFR                  Q49162 methanosarc
   14         7      1.2     255  1  UNG_HSVE2                   P53765 equine herp
   15         7      1.2     314  1  GBLP_SCHPO                  Q10281 schizosacch
   16         7      1.2     326  1  GBB2_BOVIN                  P11017 bos taurus
   17         7      1.2     331  1  YB94_METTH                  O27262 methanobact
   18         7      1.2     332  1  GBLP_DICDI                  P46800 dictyosteli
   19         7      1.2     334  1  YE10_SYNY3                  P73595 synechocyst
   20         7      1.2     340  1  GBB1_CAEEL                  P17343 caenorhabdi
   21         7      1.2     340  1  GBB1_HOMAM                  O45040 homarus ame
   22         7      1.2     340  1  GBB1_HUMAN                  P04901 homo sapien
   23         7      1.2     340  1  GBB1_RAT                    P54311 rattus norv
   24         7      1.2     340  1  GBB1_XENLA                  P79959 xenopus lae
   25         7      1.2     340  1  GBB2_HUMAN                  P11016 homo sapien
   26         7      1.2     340  1  GBB2_MOUSE                  P54312 mus musculu
   27         7      1.2     340  1  GBB2_RAT                    P54313 rattus norv
   28         7      1.2     340  1  GBB3_CANFA                  P79147 canis famil
   29         7      1.2     340  1  GBB3_HUMAN                  P16520 homo sapien
   30         7      1.2     340  1  GBB3_MOUSE                  Q61011 mus musculu
   31         7      1.2     340  1  GBB3_RAT                    P52287 rattus norv
   32         7      1.2     340  1  GBB4_MOUSE                  P29387 mus musculu
   33         7      1.2     341  1  GBB_LYMST                   Q08706 lymnaea sta
```

```
   34         7      1.2     371  1  YM18_PSEAE                  Q01609 pseudomonas
   35         7      1.2     394  1  BENE_ACICA                  P07775 acinetobact
   36         7      1.2     415  1  SCR_DROME                   P09077 drosophila
   37         7      1.2     442  1  SYH_TREPA                   O83647 treponema p
   38         7      1.2     448  1  TRN4_ECOLI                  Q60216 escherichia
   39         7      1.2     488  1  VPAP_HSV11                  P10226 herpes simp
   40         7      1.2     494  1  PAC1_YEAST                  P39946 saccharomyc
   41         7      1.2     522  1  COR1_SCHPO                  O13688 schizosacch
   42         7      1.2     533  1  DHAX_YEAST                  P22281 saccharomyc
   43         7      1.2     541  1  GATB_YEAST                  P33893 saccharomyc
   44         7      1.2     576  1  PWP1_YEAST                  P21304 saccharomyc
   45         7      1.2     604  1  RCO1_NEUCR                  P78706 neurospora

                                 ALIGNMENTS

RESULT   1
SE10_CAEEL
ID   SE10_CAEEL     STANDARD;      PRT;   579 AA.
AC   Q93794;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   SEL-10 PROTEIN.
GN   SEL-10 OR F55B12.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Sims M.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   REVISIONS.
RC   STRAIN=BRISTOL N2;
RA   Jones S.J.M.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z79757; CAB02129.1; -.
DR   WormPep; F55B12.3; CE16120.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 5.
DR   PROSITE; PS50082; WD_REPEATS_2; 7.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   DOMAIN      113    159       F-BOX.
FT   REPEAT      245    274       WD 1.
FT   REPEAT      286    316       WD 2.
FT   REPEAT      328    356       WD 3.
FT   REPEAT      368    396       WD 4.
FT   REPEAT      408    438       WD 5.
FT   REPEAT      453    481       WD 6.
FT   REPEAT      493    522       WD 7.
SQ   SEQUENCE   579 AA;  64275 MW;  0647245AF5964663 CRC64;
```

```
  Query Match             2.0%;  Score 12;  DB 1;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 0.0006;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     438 SGSLDTSIRVWD 449
           ||||||||||||
Db     427 SGSLDTSIRVWD 438


RESULT   2
GROG_BOVIN
ID   GROG_BOVIN     STANDARD;      PRT;    98 AA.
AC   O46675;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   GROWTH REGULATED PROTEIN HOMOLOG GAMMA PRECURSOR (GRO-GAMMA).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yoshimura T., Modi W.S.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC       C-X-C) (CHEMOKINE CXC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U95811; AAB93927.1; -.
DR   InterPro; IPR001089; -.
DR   InterPro; IPR001811; -.
DR   Pfam; PF00048; IL8; 1.
DR   PRINTS; PR00437; SMALLCYTKCXC.
DR   PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW   Cytokine; Growth factor; Inflammatory response; Signal.
FT   SIGNAL        1     29       POTENTIAL.
FT   CHAIN        30     98       GROWTH REGULATED PROTEIN HOMOLOG GAMMA.
FT   DISULFID     39     65       BY SIMILARITY.
FT   DISULFID     41     81       BY SIMILARITY.
SQ   SEQUENCE   98 AA;  10393 MW;  942CD6897C21EDE9 CRC64;


  Query Match             1.4%;  Score 8;  DB 1;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     494 QCLQTLQG 501
           ||||||||
Db      40 QCLQTLQG 47


RESULT   3
GRO2_RABIT
ID   GRO2_RABIT     STANDARD;      PRT;   104 AA.
AC   P47854;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   GROWTH REGULATED PROTEIN HOMOLOG PRECURSOR (GRO HOMOLOG).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schwartz D., Chaverri-Alamada L., Berliner J., Kirchgessner T.,
RA   Quisomoro D., Fang J., Tekamp-Olson P., Lusis J., Fogelman A.,
RA   Territo M.;
RL   Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A ROLE IN MONOCYTE ADHESION TO THE ENDOTHELIUM.
CC   -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC       C-X-C) (CHEMOKINE CXC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U12310; AAA20487.1; -.
DR   HSSP; P09341; 1MGS.
DR   InterPro; IPR001089; -.
DR   InterPro; IPR001811; -.
DR   InterPro; IPR002473; -.
DR   Pfam; PF00048; IL8; 1.
DR   PRINTS; PR00436; INTERLEUKIN8.
DR   PRINTS; PR00437; SMALLCYTKCXC.
DR   PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW   Cytokine; Growth factor; Inflammatory response; Signal.
FT   SIGNAL        1     31       PROBABLE.
FT   CHAIN        32    104       GROWTH REGULATED PROTEIN HOMOLOG.
FT   DISULFID     40     66       BY SIMILARITY.
FT   DISULFID     42     82       BY SIMILARITY.
SQ   SEQUENCE   104 AA;  10900 MW;  10B9D07B65C77F67 CRC64;


  Query Match             1.4%;  Score 8;  DB 1;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     494 QCLQTLQG 501
           ||||||||
Db      41 QCLQTLQG 48


RESULT   4
GROA_BOVIN
ID   GROA_BOVIN     STANDARD;      PRT;   104 AA.
AC   O46676;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   GROWTH REGULATED PROTEIN HOMOLOG ALPHA PRECURSOR (GRO-ALPHA).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yoshimura T., Modi W.S.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC       C-X-C) (CHEMOKINE CXC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U95812; AAB93928.1; -.
```

```
DR   InterPro; IPR001089; -.
DR   InterPro; IPR001811; -.
DR   InterPro; IPR002473; -.
DR   Pfam; PF00048; IL8; 1.
DR   PRINTS; PR00436; INTERLEUKIN8.
DR   PRINTS; PR00437; SMALLCYTKCXC.
DR   PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW   Cytokine; Growth factor; Inflammatory response; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    104       GROWTH REGULATED PROTEIN HOMOLOG ALPHA.
FT   DISULFID     40     66       BY SIMILARITY.
FT   DISULFID     42     82       BY SIMILARITY.
SQ   SEQUENCE   104 AA;  10984 MW;  1002CAC064DB1F76 CRC64;


  Query Match             1.4%;  Score 8;  DB 1;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     41 QCLQTLQG 48


RESULT   5
GROB_BOVIN
ID   GROB_BOVIN     STANDARD;      PRT;   104 AA.
AC   O46677;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   GROWTH REGULATED PROTEIN HOMOLOG BETA PRECURSOR (GRO-BETA).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yoshimura T., Modi W.S.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC       C-X-C) (CHEMOKINE CXC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U95813; AAB93929.1; -.
DR   InterPro; IPR001089; -.
DR   InterPro; IPR001811; -.
DR   InterPro; IPR002473; -.
DR   Pfam; PF00048; IL8; 1.
DR   PRINTS; PR00436; INTERLEUKIN8.
DR   PRINTS; PR00437; SMALLCYTKCXC.
DR   PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW   Cytokine; Growth factor; Inflammatory response; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    104       GROWTH REGULATED PROTEIN HOMOLOG BETA.
FT   DISULFID     40     66       BY SIMILARITY.
FT   DISULFID     42     82       BY SIMILARITY.
SQ   SEQUENCE   104 AA;  10950 MW;  40A8C06A64D67F7B CRC64;


  Query Match             1.4%;  Score 8;  DB 1;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     41 QCLQTLQG 48


RESULT   6
GRO_HUMAN
ID   GRO_HUMAN      STANDARD;      PRT;   107 AA.
AC   P09341;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   GROWTH REGULATED PROTEIN PRECURSOR (MELANOMA GROWTH STIMULATORY
DE   ACTIVITY) (MGSA) (NEUTROPHIL-ACTIVATING PROTEIN 3) (NAP-3).
GN   SCYB1 OR GRO1 OR GROA OR GRO OR MGSA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88041072; PubMed=2890161;
RA   Anisowicz A., Bardwell L., Sager R.;
RT   "Constitutive overexpression of a growth-regulated gene in
RT   transformed Chinese hamster and human cells.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:7188-7192(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88328991; PubMed=2970963;
RA   Richmond A., Balentien E., Thomas H.G., Flaggs G., Barton D.E.,
RA   Spiess J., Bordoni R., Francke U., Derynck R.;
RT   "Molecular characterization and chromosomal mapping of melanoma
RT   growth stimulatory activity, a growth factor structurally related to
RT   beta-thromboglobulin.";
RL   EMBO J. 7:2025-2033(1988).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=91057157; PubMed=2129556;
RA   Baker N.E., Kucera G., Richmond A.;
RT   "Nucleotide sequence of the human melanoma growth stimulatory
RT   activity (MGSA) gene.";
RL   Nucleic Acids Res. 18:6453-6453(1990).
RN   [4]
RP   SEQUENCE OF 35-65.
RX   MEDLINE=90217938; PubMed=2182761;
RA   Schroeder J.-M., Persoon N.L.M., Christophers E.;
RT   "Lipopolysaccharide-stimulated human monocytes secrete, apart from
RT   neutrophil-activating peptide 1/interleukin 8, a second neutrophil-
RT   activating protein. NH2-terminal amino acid sequence identity with
RT   melanoma growth stimulatory activity.";
RL   J. Exp. Med. 171:1091-1100(1990).
RN   [5]
RP   SEQUENCE OF 35-57.
RX   MEDLINE=89246368; PubMed=2655583;
RA   Golds E.E., Mason P., Nyirkos P.;
RT   "Inflammatory cytokines induce synthesis and secretion of gro protein
RT   and a neutrophil chemotactic factor but not beta 2-microglobulin in
RT   human synovial cells and fibroblasts.";
RL   Biochem. J. 259:585-588(1989).
RN   [6]
RP   POSSIBLE FUNCTION.
RX   MEDLINE=89356650; PubMed=2670560;
RA   Wen D., Rowland A., Derynck R.;
RT   "Expression and secretion of gro/MGSA by stimulated human endothelial
RT   cells.";
RL   EMBO J. 8:1761-1766(1989).
RN   [7]
RP   STRUCTURE BY NMR.
RX   MEDLINE=93387459; PubMed=8397104;
RA   Fairbrother W.J., Reilly D., Colby T., Horuk R.;
RT   "1H assignment and secondary structure determination of human
RT   melanoma growth stimulating activity (MGSA) by NMR spectroscopy.";
RL   FEBS Lett. 330:302-306(1993).
```

```
RN   [8]
RP   STRUCTURE BY NMR.
RX   MEDLINE=94376296; PubMed=8089846;
RA   Fairbrother W.J., Reilly D., Colby T., Hesselgesser J., Horuk R.;
RT   "The solution structure of melanoma growth stimulating activity.";
RL   J. Mol. Biol. 242:252-270(1994).
RN   [9]
RP   STRUCTURE BY NMR.
RX   MEDLINE=95105175; PubMed=7806518;
RA   Kim K.S., Clark-Lewis I., Sykes B.D.;
RT   "Solution structure of GRO/melanoma growth stimulatory activity
RT   determined by 1H NMR spectroscopy.";
RL   J. Biol. Chem. 269:32909-32915(1994).
CC   -!- FUNCTION: HAS CHEMOTACTIC ACTIVITY FOR NEUTROPHILS. MAY PLAY A
CC       ROLE IN INFLAMMATION AND EXERTS ITS EFFECTS ON ENDOTHELIAL CELLS
CC       IN AN AUTOCRINE FASHION.
CC   -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC       C-X-C) (CHEMOKINE CXC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J03561; AAA35933.1; -.
DR   EMBL; X12510; CAA31027.1; -.
DR   EMBL; X54489; CAA38361.1; -.
DR   PIR; A28414; A28414.
DR   PIR; S00983; S00983.
DR   PIR; S03976; S03976.
DR   PIR; S13669; S13669.
DR   PDB; 1MGS; 30-SEP-94.
DR   PDB; 1MSG; 31-MAR-95.
DR   PDB; 1MSH; 31-MAR-95.
DR   MIM; 155730; -.
DR   InterPro; IPR001089; -.
DR   InterPro; IPR001811; -.
DR   InterPro; IPR002473; -.
DR   Pfam; PF00048; IL8; 1.
DR   PRINTS; PR00436; INTERLEUKIN8.
DR   PRINTS; PR00437; SMALLCYTKCXC.
DR   PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW   Cytokine; Growth factor; Inflammatory response; Signal; 3D-structure.
FT   SIGNAL        1     34
FT   CHAIN        35    107       GROWTH REGULATED PROTEIN.
FT   DISULFID     43     69
FT   DISULFID     45     85
SQ   SEQUENCE   107 AA;  11301 MW;  17048A6B4D765CA2 CRC64;


  Query Match             1.4%;  Score 8;  DB 1;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    494 QCLQTLQG 501
          ||||||||
Db     44 QCLQTLQG 51


RESULT  7
MI2A_HUMAN
ID   MI2A_HUMAN     STANDARD;      PRT;   107 AA.
AC   P19875; Q9UPB8;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MACROPHAGE INFLAMMATORY PROTEIN-2-ALPHA PRECURSOR (MIP2-ALPHA) (GROWTH
DE   REGULATED PROTEIN BETA) (GRO-BETA).
GN   GRO2 OR GROB OR MIP2A.
OS   Homo sapiens (Human).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Histiocytic lymphoma;
RX   MEDLINE=90354792; PubMed=2201751;
RA   Tekamp-Olson P., Gallegos C., Bauer D., McClain J., Sherry B.,
RA   Fabre M., van Deventer S., Cerami A.;
RT   "Cloning and characterization of cDNAs for murine macrophage
RT   inflammatory protein 2 and its human homologues.";
RL   J. Exp. Med. 172:911-919(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90377259; PubMed=2078213;
RA   Iida N., Grotendorst G.R.;
RT   "Cloning and sequencing of a new gro transcript from activated human
RT   monocytes: expression in leukocytes and wound tissue.";
RL   Mol. Cell. Biol. 10:5596-5599(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91017578; PubMed=2217207;
RA   Haskill S., Peace A., Morris J., Sporn S.A., Anisowicz A., Lee S.W.,
RA   Smith T., Martin G., Ralph P., Sager R.;
RT   "Identification of three related human GRO genes encoding cytokine
RT   functions.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:7732-7736(1990).
RN   [4]
RP   SEQUENCE OF 35-107 FROM N.A.
RA   Jang J.S., Kim B.E.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   STRUCTURE BY NMR OF 39-107.
RX   MEDLINE=20069929; PubMed=10600366;
RA   Qian Y.Q., Johanson K.O., McDevitt P.;
RT   "Nuclear magnetic resonance solution structure of truncated human
RT   GRObeta [5-73] and its structural comparison with CXC chemokine
RT   family members GROalpha and IL-8.";
RL   J. Mol. Biol. 294:1065-1072(1999).
CC   -!- FUNCTION: PRODUCED BY ACTIVATED MONOCYTES AND NEUTROPHILS AND
CC       EXPRESSED AT SITES OF INFLAMMATION.
CC   -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC       C-X-C) (CHEMOKINE CXC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X53799; CAA37808.1; -.
DR   EMBL; M36820; AAA63183.1; -.
DR   EMBL; M57731; AAA63182.1; -.
DR   EMBL; AF043340; AAC03540.1; -.
DR   PIR; JH0281; JH0281.
DR   PDB; 1QNK; 04-FEB-00.
DR   MIM; 139110; -.
DR   InterPro; IPR001089; -.
DR   InterPro; IPR001811; -.
DR   InterPro; IPR002473; -.
DR   Pfam; PF00048; IL8; 1.
DR   PRINTS; PR00436; INTERLEUKIN8.
DR   PRINTS; PR00437; SMALLCYTKCXC.
DR   PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW   Cytokine; Chemotaxis; Inflammatory response; Signal; 3D-structure.
FT   SIGNAL        1     34
FT   CHAIN        35    107       MACROPHAGE INFLAMMATORY PROTEIN-
FT                                2-ALPHA.
FT   DISULFID     43     69
FT   DISULFID     45     85
SQ   SEQUENCE   107 AA;  11389 MW;  740F277E928571BA CRC64;
```

```
  Query Match             1.4%;  Score 8;  DB 1;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     494 QCLQTLQG 501
           ||||||||
Db      44 QCLQTLQG 51


RESULT  8
MI2B_HUMAN
ID   MI2B_HUMAN     STANDARD;      PRT;   107 AA.
AC   P19876;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MACROPHAGE INFLAMMATORY PROTEIN-2-BETA PRECURSOR (MIP2-BETA) (GROWTH
DE   REGULATED PROTEIN GAMMA) (GRO-GAMMA).
GN   GRO3 OR GROG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Histiocytic lymphoma;
RX   MEDLINE=90354792; PubMed=2201751;
RA   Tekamp-Olson P., Gallegos C., Bauer D., McClain J., Sherry B.,
RA   Fabre M., van Deventer S., Cerami A.;
RT   "Cloning and characterization of cDNAs for murine macrophage
RT   inflammatory protein 2 and its human homologues.";
RL   J. Exp. Med. 172:911-919(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91017578; PubMed=2217207;
RA   Haskill S., Peace A., Morris J., Sporn S.A., Anisowicz A., Lee S.W.,
RA   Smith T., Martin G., Ralph P., Sager R.;
RT   "Identification of three related human GRO genes encoding cytokine
RT   functions.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:7732-7736(1990).
CC   -!- FUNCTION: MAY PLAY A ROLE IN INFLAMMATION AND EXERT ITS EFFECTS
CC       ON ENDOTHELIAL CELLS IN AN AUTOCRINE FASHION.
CC   -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC       C-X-C) (CHEMOKINE CXC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X53800; CAA37809.1; -.
DR   EMBL; M36821; AAA63184.1; -.
DR   PIR; B38290; B38290.
DR   PIR; JH0282; JH0282.
DR   HSSP; P09341; 1MGS.
DR   MIM; 139111; -.
DR   InterPro; IPR001089; -.
DR   InterPro; IPR001811; -.
DR   InterPro; IPR002473; -.
DR   Pfam; PF00048; IL8; 1.
DR   PRINTS; PR00436; INTERLEUKIN8.
DR   PRINTS; PR00437; SMALLCYTKCXC.
DR   PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW   Cytokine; Chemotaxis; Inflammatory response; Signal.
FT   SIGNAL        1     34
FT   CHAIN        35    107       MACROPHAGE INFLAMMATORY PROTEIN-
FT                                2-BETA.
FT   DISULFID     43     69       BY SIMILARITY.
FT   DISULFID     45     85       BY SIMILARITY.
FT   CONFLICT     27     28       AA -> G (IN REF. 2).
SQ   SEQUENCE   107 AA;  11342 MW;  97A69946B7F1F070 CRC64;

  Query Match             1.4%;  Score 8;  DB 1;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     494 QCLQTLQG 501
           ||||||||
Db      44 QCLQTLQG 51


RESULT  9
DRRA_STRPE
ID   DRRA_STRPE     STANDARD;      PRT;   330 AA.
AC   P32010;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   DAUNORUBICIN RESISTANCE ATP-BINDING PROTEIN DRRA.
GN   DRRA.
OS   Streptomyces peucetius.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1950;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29050;
RX   MEDLINE=92020891; PubMed=1924314;
RA   Guilfoile P.G., Hutchinson C.R.;
RT   "A bacterial analog of the mdr gene of mammalian tumor cells is
RT   present in Streptomyces peucetius, the producer of daunorubicin and
RT   doxorubicin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:8553-8557(1991).
CC   -!- FUNCTION: DRRA AND DRRB MAY ACT JOINTLY TO CONFER DAUNORUBICIN AND
CC       DOXORUBICIN RESISTANCE BY AN EXPORT MECHANISM.
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M73758; AAA74717.1; -.
DR   PIR; S27707; S27707.
DR   InterPro; IPR001617; -.
DR   Pfam; PF00005; ABC_tran; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Transport; Antibiotic resistance.
FT   NP_BIND      41     48       ATP (BY SIMILARITY).
SQ   SEQUENCE   330 AA;  35700 MW;  582D66C90D54E6B9 CRC64;


  Query Match             1.4%;  Score 8;  DB 1;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 4;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      59 EVRSFSLG 66
           ||||||||
Db     292 EVRSFSLG 299


RESULT  10
YSS1_CAEEL
ID   YSS1_CAEEL     STANDARD;      PRT;   701 AA.
AC   Q09990;
DT   15-JUL-1998 (Rel. 36, Created)
```

```
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   HYPOTHETICAL 80.3 KDA TRP-ASP REPEATS CONTAINING PROTEIN K10B2.1 IN
DE   CHROMOSOME II.
GN   K10B2.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Miller N.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   -!- SIMILARITY: STRONG, TO X.LAEVIS BETA-TRCP.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U28730; AAA68258.1; -.
DR   WormPep; K10B2.1; CE02008.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 5.
DR   PROSITE; PS50082; WD_REPEATS_2; 7.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Hypothetical protein; Repeat; WD repeat.
FT   REPEAT      256    284       WD 1.
FT   REPEAT      296    324       WD 2.
FT   REPEAT      336    364       WD 3.
FT   REPEAT      379    407       WD 4.
FT   REPEAT      419    447       WD 5.
FT   REPEAT      459    487       WD 6.
FT   REPEAT      508    536       WD 7.
FT   DOMAIN      606    615       POLY-ALA.
SQ   SEQUENCE   701 AA;  80320 MW;  69FA0B00F83270E3 CRC64;


  Query Match             1.4%;  Score 8;  DB 1;  Length 701;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     446 RVWDVETG 453
           ||||||||
Db     321 RVWDVETG 328


RESULT  11
CC4H_HUMAN
ID   CC4H_HUMAN     STANDARD;      PRT;  1918 AA.
AC   P50851;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CDC4-LIKE PROTEIN (BEIGE-LIKE PROTEIN) (FRAGMENT).
GN   CDC4L OR BGL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mager D.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
```

```
RP   SEQUENCE OF 1405-1918 FROM N.A.
RX   MEDLINE=92372019; PubMed=1505956;
RA   Feuchter A.E., Freeman J.D., Mager D.L.;
RT   "Strategy for detecting cellular transcripts promoted by human
RT   endogenous long terminal repeats: identification of a novel gene
RT   (CDC4L) with homology to yeast CDC4.";
RL   Genomics 13:1237-1246(1992).
CC   -!- SIMILARITY: CONTAINS AT LEAST 2 WD REPEATS (TRP-ASP DOMAINS).
CC   -!- SIMILARITY: CONTAINS 1 BEACH DOMAIN.
CC   -!- CAUTION: WAS SAID TO BE SIMILAR TO YEAST CDC4, BUT THAT SIMILARITY
CC       IS VERY LIMITED.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M83822; AAB09603.1; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 2.
DR   PROSITE; PS50197; BEACH; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
DR   PROSITE; PS50082; WD_REPEATS_2; FALSE_NEG.
DR   PROSITE; PS50294; WD_REPEATS_REGION; FALSE_NEG.
KW   Repeat; WD repeat.
FT   NON_TER       1      1
FT   DOMAIN     1445   1715       BEACH.
FT   REPEAT     1817   1859       WD 2.
FT   REPEAT     1862   1905       WD 3.
SQ   SEQUENCE   1918 AA;  213598 MW;  427EA619BDF9A724 CRC64;


  Query Match             1.4%;  Score 8;  DB 1;  Length 1918;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     358 SGSRDATL 365
           ||||||||
Db    1885 SGSRDATL 1892


RESULT  12
FER5_MAIZE
ID   FER5_MAIZE     STANDARD;      PRT;   135 AA.
AC   P27789;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   FERREDOXIN V PRECURSOR (FD V).
GN   FDX5 OR PFD5.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC   Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hase T., Kimatsa Y., Yonekura K., Matsumura T., Sakakibara H.;
RT   "Molecular cloning and differential expression of the maize ferredoxin
RT   gene family.";
RL   Plant Physiol. 96:77-83(1991).
CC   -!- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC       ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
CC   -!- COFACTOR: BINDS A 2FE-2S CLUSTER.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M73828; AAA33462.1; -.
DR   HSSP; P27320; 1DOY.
DR   MaizeDB; 66392; -.
DR   InterPro; IPR000564; -.
DR   InterPro; IPR001041; -.
DR   Pfam; PF00111; fer2; 1.
DR   PRINTS; PR00159; 2FE2SFRDOXIN.
DR   PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
KW   Electron transport; Iron-sulfur; Chloroplast; Transit peptide;
KW   Multigene family.
FT   TRANSIT       1     38       CHLOROPLAST (BY SIMILARITY).
FT   CHAIN        39    135       FERREDOXIN V.
FT   METAL        77     77       IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT   METAL        82     82       IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT   METAL        85     85       IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT   METAL       115    115       IRON-SULFUR (2FE-2S) (BY SIMILARITY).
SQ   SEQUENCE   135 AA;  14399 MW;  8FA43C41AD4CB976 CRC64;

  Query Match              1.2%;  Score 7;  DB 1;  Length 135;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    436 VVSGSLD 442
          |||||||
Db     89 VVSGSLD 95


RESULT  13
DCB1_METFR
ID   DCB1_METFR     STANDARD;      PRT;   170 AA.
AC   Q49162;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE I BETA SUBUNIT (EC 1.2.99.2).
GN   CDHIB.
OS   Methanosarcina frisia.
OC   Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;
OC   Methanosarcina.
OX   NCBI_TaxID=2212;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GOE1;
RX   MEDLINE=96278885; PubMed=8662887;
RA   Eggen R.I.L., van Kranenburg R., Vriesema A.J.M., Geerling A.C.M.,
RA   Verhagen M.F.J.M., Hagen W.R., de Vos W.M.;
RT   "Carbon monoxide dehydrogenase from Methanosarcina frisia Gol.
RT   Characterization of the enzyme and the regulated expression of two
RT   operon-like cdh gene clusters.";
RL   J. Biol. Chem. 271:14256-14263(1996).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: NICKEL ION.
CC   -!- SUBUNIT: HOMOTETRAMER OF TWO ALPHA AND TWO BETA CHAINS.
CC   -!- MISCELLANEOUS: OPTIMUM ACTIVITY IS MEASURED BETWEEN PH 8 AND 9.
CC   -!- SIMILARITY: TO M.SOEHNGENII CARBON MONOXIDE DEHYDROGENASE.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L26487; AAC37045.1; -.
KW   Oxidoreductase; Nickel.
SQ   SEQUENCE   170 AA;  18696 MW;  74B309E6487E3E79 CRC64;
```

```
  Query Match              1.2%;  Score 7;  DB 1;  Length 170;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    134 ALDELID 140
          |||||||
Db    162 ALDELID 168


RESULT  14
UNG_HSVE2
ID   UNG_HSVE2      STANDARD;      PRT;   255 AA.
AC   P53765;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   URACIL-DNA GLYCOSYLASE (EC 3.2.2.-) (UDG).
GN   46.
OS   Equine herpesvirus type 2 (strain 86/87) (EHV-2).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae.
OX   NCBI_TaxID=82831;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95302501; PubMed=7783207;
RA   Telford E.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;
RT   "The DNA sequence of equine herpesvirus 2.";
RL   J. Mol. Biol. 249:520-528(1995).
CC   -!- FUNCTION: EXCISES URACIL RESIDUES FROM THE DNA WHICH CAN ARISE
CC       AS A RESULT OF MISINCORPORATION OF DUMP RESIDUES BY DNA
CC       POLYMERASE OR DUE TO DEAMINATION OF CYTOSINE.
CC   -!- SIMILARITY: BELONGS TO THE URACIL-DNA GLYCOSYLASE FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U20824; AAC13834.1; -.
DR   HSSP; P13051; 1AKZ.
DR   InterPro; IPR002043; -.
DR   Pfam; PF00315; UNG; 1.
DR   PROSITE; PS00130; U_DNA_GLYCOSYLASE; 1.
KW   DNA repair; Hydrolase; Glycosidase.
FT   ACT_SITE     90     90       GENERAL BASE (BY SIMILARITY).
SQ   SEQUENCE   255 AA;  29099 MW;  20104402C5297336 CRC64;

  Query Match              1.2%;  Score 7;  DB 1;  Length 255;
  Best Local Similarity   100.0%;  Pred. No: 33;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    142 CEPTQVK 148
          |||||||
Db     77 CEPTQVK 83


RESULT  15
GBLP_SCHPO
ID   GBLP_SCHPO     STANDARD;      PRT;   314 AA.
AC   Q10281; P78896;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT-LIKE PROTEIN.
GN   SPAC6B12.15.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
```

```
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ED616;
RA   Park S.K., Yoo H.S.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 5-314 FROM N.A.
RC   STRAIN=PR745;
RA   Yoshioka S., Kato K., Okayama H.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L37885; AAA56865.1; -.
DR   EMBL; Z98531; CAB11079.1; -.
DR   EMBL; D89247; BAA13908.1; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00678; WD_REPEATS_1; 4.
DR   PROSITE; PS50082; WD_REPEATS_2; 6.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   REPEAT        13     44       WD 1.
FT   REPEAT        61     91       WD 2.
FT   REPEAT       103    133       WD 3.
FT   REPEAT       146    178       WD 4.
FT   REPEAT       190    220       WD 5.
FT   REPEAT       231    260       WD 6.
FT   REPEAT       281    311       WD 7.
FT   CONFLICT       5      5       L -> I (IN REF. 2).
FT   CONFLICT       7      7       L -> V (IN REF. 2).
FT   CONFLICT      89     89       L -> W (IN REF. 2).
FT   CONFLICT     137    137       N -> Y (IN REF. 2).
FT   CONFLICT     168    168       A -> V (IN REF. 2).
FT   CONFLICT     171    171       D -> E (IN REF. 2).
FT   CONFLICT     173    173       A -> P (IN REF. 2).
FT   CONFLICT     176    176       V -> A (IN REF. 2).
FT   CONFLICT     244    244       N -> D (IN REF. 2).
SQ   SEQUENCE   314 AA;  34851 MW;  4E14707164E68ACD CRC64;

  Query Match             1.2%;  Score 7;  DB 1;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        356 VVSGSRD 362
              |||||||
Db        120 VVSGSRD 126


Search completed: September 25, 2001, 14:49:27
Job time: 206 sec
```

```
                      GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         September 25, 2001, 14:45:36 ; Search time 22.42 Seconds
                                               (without alignments)
                                               3475.810 Million cell updates/sec

Title:          US-09-328-877A-8
Perfect score:  589
Sequence:       1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:  OLIGO
                Gapop 60.0 , Gapext 60.0

Searched:       425026 seqs, 132305027 residues

Word size :     0

Total number of hits satisfying chosen parameters:      425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :      SPTREMBL_16:*
               1:  sp_archea:*
               2:  sp_bacteria:*
               3:  sp_fungi:*
               4:  sp_human:*
               5:  sp_invertebrate:*
               6:  sp_mammal:*
               7:  sp_mhc:*
               8:  sp_organelle:*
               9:  sp_phage:*
               10:  sp_plant:*
               11:  sp_rodent:*
               12:  sp_unclassified:*
               13:  sp_vertebrate:*
               14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                     %
Result             Query
  No.    Score   Match Length DB  ID                         Description
-----------------------------------------------------------------------------
    1      540    91.7    553  4  Q9NUX6                     Q9nux6 homo sapien
    2       21     3.6   1326  5  Q9VZF4                     Q9vzf4 drosophila
    3       12     2.0    587  5  O44083                     O44083 caenorhabdi
    4       10     1.7    310  4  O75938                     O75938 homo sapien
    5       10     1.7    357  4  O95320                     O95320 homo sapien
    6        9     1.5    472  4  Q9Y6J1                     Q9y6j1 homo sapien
    7        9     1.5    524  4  Q9NSU1                     Q9nsu1 homo sapien
    8        9     1.5    703  11  Q9WVM9                     Q9wvm9 mus musculu
    9        9     1.5    741  4  Q9H9K3                     Q9h9k3 homo sapien
   10        9     1.5    845  10  Q9LJR3                     Q9ljr3 arabidopsis
   11        9     1.5   1028  4  Q9HCM8                     Q9hcm8 homo sapien
   12        9     1.5   2904  11  Q9EPN0                     Q9epn0 mus musculu
   13        9     1.5   2931  11  Q9EPM9                     Q9epm9 mus musculu
   14        9     1.5   2936  11  Q9EPN1                     Q9epn1 mus musculu
   15        8     1.4    272  2  O66716                     O66716 aquifex aeo
   16        8     1.4    276  5  Q9VXP8                     Q9vxp8 drosophila
   17        8     1.4    326  5  Q9XYF7                     Q9xyf7 drosophila
   18        8     1.4    326  5  Q9VAJ2                     Q9vaj2 drosophila
   19        8     1.4    327  5  O76523                     O76523 drosophila
```

```
   20        8     1.4    337  4  Q9H2U3                     Q9h2u3 homo sapien
   21        8     1.4    344  13  Q9PSJ1                     Q9psj1 ictalurus p
   22        8     1.4    386  5  Q9XWU3                     Q9xwu3 caenorhabdi
   23        8     1.4    476  5  Q9VGI7                     Q9vgi7 drosophila
   24        8     1.4    507  5  Q09661                     Q09661 caenorhabdi
   25        8     1.4    543  5  Q9V6J3                     Q9v6j3 drosophila
   26        8     1.4    665  5  Q9GNN6                     Q9gnn6 caenorhabdi
   27        8     1.4    821  10  O04093                     O04093 arabidopsis
   28        8     1.4   2792  11  Q9ESD4                     Q9esd4 mus musculu
   29        8     1.4   2856  11  Q9ESE1                     Q9ese1 mus musculu
   30        7     1.2     41  14  Q83070                     Q83070 leishmania
   31        7     1.2     41  14  Q83098                     Q83098 leishmania
   32        7     1.2     44  14  Q9QEI2                     Q9qei2 human immun
   33        7     1.2     52  2  O33534                     O33534 rhizobium l
   34        7     1.2    123  11  Q61621                     Q61621 mus musculu
   35        7     1.2    139  2  Q46367                     Q46367 chlorobium
   36        7     1.2    139  2  Q46466                     Q46466 chlorobium
   37        7     1.2    149  5  Q21091                     Q21091 caenorhabdi
   38        7     1.2    150  2  Q9WWX4                     Q9wwx4 pseudomonas
   39        7     1.2    150  5  Q06784                     Q06784 haematobia
   40        7     1.2    163  2  Q9X5U1                     Q9x5u1 streptomyce
   41        7     1.2    164  5  Q9TYL1                     Q9tyl1 caenorhabdi
   42        7     1.2    170  1  P72020                     P72020 methanosarc
   43        7     1.2    176  5  Q9GV52                     Q9gv52 drosophila
   44        7     1.2    187  1  Q9YF07                     Q9yf07 aeropyrum p
   45        7     1.2    190  2  O86782                     O86782 streptomyce

                                  ALIGNMENTS


RESULT   1
Q9NUX6
ID   Q9NUX6      PRELIMINARY;      PRT;   553 AA.
AC   Q9NUX6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CDNA FLJ11071 FIS, CLONE PLACE1004937, MODERATELY SIMILAR TO SEL-10
DE   PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA   Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA   Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK001933; BAA91986.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
DR   SMART; SM00256; FBOX; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   553 AA;  62280 MW;  CA829C221986A3F2 CRC64;


  Query Match             91.7%;  Score 540;  DB 4;  Length 553;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  540;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         50 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 MKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITS 73
```

```
Qy     110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 133

Qy     170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 229
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFI 193

Qy     230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 289
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194 HSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWS 253

Qy     290 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 349
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     254 AVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCM 313

Qy     350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 409
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     314 HLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWD 373

Qy     410 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     374 PETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGM 433

Qy     470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     434 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 493

Qy     530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     494 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 553


RESULT   2
Q9VZF4
ID   Q9VZF4      PRELIMINARY;      PRT;  1326 AA.
AC   Q9VZF4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CG15010 PROTEIN.
GN   CG15010.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
```

```
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003480; AAG22246.1; -.
DR   FlyBase; FBgn0035516; CG15010.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; 5.
KW   Repeat; WD repeat.
SQ   SEQUENCE   1326 AA;  141360 MW;  3F42C873CFA3027F CRC64;


  Query Match             3.6%;  Score 21;  DB 5;  Length 1326;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-12;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     435 HVVSGSLDTSIRVWDVETGNC 455
           |||||||||||||||||||||
Db    1167 HVVSGSLDTSIRVWDVETGNC 1187


RESULT   3
O44083
ID   O44083      PRELIMINARY;      PRT;   587 AA.
AC   O44083;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   SEL-10.
GN   SEL-10.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98051191; PubMed=9389650;
RA   Hubbard E.J.A., Wu G., Kitajewski J., Greenwald I.;
RT   "sel-10, a negative regulator of lin-12 activity in Caenorhabditis
RT   elegans, encodes a member of the CDC4 family of proteins.";
RL   Genes Dev. 11:3182-3193(1997).
DR   EMBL; AF020788; AAC47809.1; -.
DR   InterPro; IPR000412; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00890; ABC2_MEMBRANE; UNKNOWN_1.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
KW   Repeat; WD repeat.
SQ   SEQUENCE   587 AA;  65311 MW;  2D3970B4EFAA1B8C CRC64;
```

```
  Query Match             2.0%;  Score 12;  DB 5;  Length 587;
  Best Local Similarity   100.0%;  Pred. No. 0.0019;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     438 SGSLDTSIRVWD 449
           ||||||||||||
Db     435 SGSLDTSIRVWD 446

RESULT   4
O75938
ID   O75938      PRELIMINARY;      PRT;   310 AA.
AC   O75938;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   38 KDA SPLICING FACTOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98400255; PubMed=9731529;
RA   Neubauer G., King A., Rappsilber J., Calvio C., Watson M., Ajuh P.,
RA   Sleeman J., Lamond A., Mann M.;
RT   "Mass spectrometry and EST-database searching allows characterization
RT   of the multi-protein spliceosome complex.";
RL   Nat. Genet. 20:46-50(1998).
DR   EMBL; AF083383; AAC64084.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR002106; -.
DR   Pfam; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
DR   SMART; SM00320; WD40; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   310 AA;  34290 MW;  7347F144E7D9796E CRC64;

  Query Match             1.7%;  Score 10;  DB 4;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     523 SDDGTVKLWD 532
           ||||||||||
Db     112 SDDGTVKLWD 121

RESULT   5
O95320
ID   O95320      PRELIMINARY;      PRT;   357 AA.
AC   O95320;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   U5 SNRNP-SPECIFIC 40 KDA PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98449962; PubMed=9774689;
RA   Achsel T., Ahrens K., Brahms H., Teigelkamp S., Luhrmann R.;
RT   "The human U5-220kD protein (hPrp8) forms a stable RNA-free complex
RT   with several U5-specific proteins, including an RNA unwindase, a
RT   homologue of ribosomal elongation factor EF-2, and a novel WD-40
RT   protein.";
RL   Mol. Cell. Biol. 18:6756-6766(1998).
DR   EMBL; AF090988; AAC69625.1; -.
DR   InterPro; IPR001680; -.
```

```
DR   InterPro; IPR002106; -.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
DR   SMART; SM00320; WD40; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   357 AA;  39298 MW;  87A265539437BD86 CRC64;

  Query Match             1.7%;  Score 10;  DB 4;  Length 357;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     523 SDDGTVKLWD 532
           ||||||||||
Db     171 SDDGTVKLWD 180

RESULT   6
Q9Y6J1
ID   Q9Y6J1      PRELIMINARY;      PRT;   472 AA.
AC   Q9Y6J1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   LYSOSOMAL TRAFFICKING REGULATOR 2 (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tchernev V.T., McMurtrie E.B., Nguyen Q.A., Mishra V.S.,
RA   Barbosa M.D.F.S., McIndoe R., Kingsmore S.F.;
RT   "Identification of LYST2, a brain-specific member of the Chediak-
RT   Higashi syndrome gene family.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF072371; AAD41633.1; -.
DR   InterPro; IPR000409; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 4.
DR   Pfam; PF02138; Beach; 1.
DR   PROSITE; PS50197; BEACH; 1.
DR   SMART; SM00320; WD40; 1.
KW   Repeat; WD repeat.
FT   NON_TER       1      1
SQ   SEQUENCE   472 AA;  52810 MW;  679A710103FDAE5E CRC64;

  Query Match             1.5%;  Score 9;  DB 4;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     357 VSGSRDATL 365
           |||||||||
Db     313 VSGSRDATL 321

RESULT   7
Q9NSU1
ID   Q9NSU1      PRELIMINARY;      PRT;   524 AA.
AC   Q9NSU1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 58.8 KDA PROTEIN.
GN   DKFZP434G0511.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Duesterhoeft A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL137748; CAB70903.1; -.
DR   InterPro; IPR000409; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 5.
DR   Pfam; PF02138; Beach; 2.
DR   PROSITE; PS50197; BEACH; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   524 AA;  58801 MW;  F4A38C3EFE58D390 CRC64;


  Query Match              1.5%;  Score 9;  DB 4;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 2.2;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     357 VSGSRDATL 365
           |||||||||
Db     318 VSGSRDATL 326


RESULT   8
Q9WVM9
ID   Q9WVM9      PRELIMINARY;      PRT;   703 AA.
AC   Q9WVM9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   LYSOSOMAL TRAFFICKING REGULATOR 2 (FRAGMENT).
GN   LYST2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tchernev V.T., McMurtrie E.B., Nguyen Q.A., Mishra V.S.,
RA   Barbosa M.D.F.S., McIndoe R., Kingsmore S.F.;
RT   "Identification of LYST2, a brain-specific member of the Chediak-
RT   Higashi syndrome gene family.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF072372; AAD41634.1; -.
DR   InterPro; IPR000409; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 5.
DR   Pfam; PF02138; Beach; 1.
DR   PROSITE; PS50197; BEACH; 1.
DR   SMART; SM00320; WD40; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   703 AA;  79349 MW;  6776B57B1D972127 CRC64;


  Query Match              1.5%;  Score 9;  DB 11;  Length 703;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     357 VSGSRDATL 365
           |||||||||
Db     511 VSGSRDATL 519


RESULT   9
Q9H9K3
ID   Q9H9K3      PRELIMINARY;      PRT;   741 AA.
AC   Q9H9K3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CDNA FLJ12686 FIS, CLONE NT2RM4002527, WEAKLY SIMILAR TO VEGETATIBLE
DE   INCOMPATIBILITY PROTEIN HET-E-1 (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK022748; BAB14222.1; -.
FT   NON_TER     741    741
SQ   SEQUENCE   741 AA;  82214 MW;  8BE2038217996D07 CRC64;


  Query Match              1.5%;  Score 9;  DB 4;  Length 741;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     521 TSSDDGTVK 529
           |||||||||
Db      84 TSSDDGTVK 92


RESULT  10
Q9LJR3
ID   Q9LJR3      PRELIMINARY;      PRT;   845 AA.
AC   Q9LJR3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PHOTOMORPHOGENESIS REPRESSOR PROTEIN-LIKE.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RA   Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   PubMed=10907853;
RA   Nakamura Y.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT   Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT   TAC and BAC clones.";
RL   DNA Res. 7:217-221(2000).
DR   EMBL; AP000413; BAB02165.1; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_2.
DR   SMART; SM00221; STYKc; 1.
KW   ATP-binding; Repeat; Transferase; WD repeat.
SQ   SEQUENCE   845 AA;  94353 MW;  EB5AADEDFD16F316 CRC64;


  Query Match              1.5%;  Score 9;  DB 10;  Length 845;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     523 SDDGTVKLW 531
```

```
           |||||||||
Db     646 SDDGTVKLW 654

RESULT  11
Q9HCM8
ID   Q9HCM8      PRELIMINARY;      PRT;  1028 AA.
AC   Q9HCM8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   KIAA1544 PROTEIN (FRAGMENT).
GN   KIAA1544.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=10997877;
RA   Nagase T., Kikuno R., Nakayama M., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes.
RT   XVIII. The complete sequences of 100 new cDNA clones from brain which
RT   code for large proteins in vitro.";
RL   DNA Res. 7:273-281(2000).
DR   EMBL; AB046764; BAB13370.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   1028 AA;  116708 MW;  1204AAA8AA03B60F CRC64;

  Query Match             1.5%;  Score 9;  DB 4;  Length 1028;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     357 VSGSRDATL 365
           |||||||||
Db     822 VSGSRDATL 830

RESULT  12
Q9EPN0
ID   Q9EPN0      PRELIMINARY;      PRT;  2904 AA.
AC   Q9EPN0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   NEUROBEACHIN.
GN   NBEA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   PubMed=11102458;
RA   Wang X., Herberg F.W., Laue M.M., Wullner C., Hu B.,
RA   Petrasch-Parwez E., Kilimann M.W.;
RT   "Neurobeachin: A protein kinase A-anchoring, beige/Chediak-Higashi
RT   protein homolog implicated in neuronal membrane traffic.";
RL   J. Neurosci. 20:8551-8565(2000).
DR   EMBL; Y18276; CAC18812.1; -.
SQ   SEQUENCE   2904 AA;  323198 MW;  B41D4ADF71AD6FF4 CRC64;

  Query Match             1.5%;  Score 9;  DB 11;  Length 2904;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     357 VSGSRDATL 365
           |||||||||
Db    2698 VSGSRDATL 2706
```

```
RESULT  13
Q9EPM9
ID   Q9EPM9      PRELIMINARY;      PRT;  2931 AA.
AC   Q9EPM9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   NEUROBEACHIN.
GN   NBEA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   PubMed=11102458;
RA   Wang X., Herberg F.W., Laue M.M., Wullner C., Hu B.,
RA   Petrasch-Parwez E., Kilimann M.W.;
RT   "Neurobeachin: A protein kinase A-anchoring, beige/Chediak-Higashi
RT   protein homolog implicated in neuronal membrane traffic.";
RL   J. Neurosci. 20:8551-8565(2000).
DR   EMBL; Y18276; CAC18813.1; -.
SQ   SEQUENCE   2931 AA;  326149 MW;  B2E2E8E87E6B2696 CRC64;

  Query Match             1.5%;  Score 9;  DB 11;  Length 2931;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     357 VSGSRDATL 365
           |||||||||
Db    2725 VSGSRDATL 2733

RESULT  14
Q9EPN1
ID   Q9EPN1      PRELIMINARY;      PRT;  2936 AA.
AC   Q9EPN1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   NEUROBEACHIN.
GN   NBEA.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   PubMed=11102458;
RA   Wang X., Herberg F.W., Laue M.M., Wullner C., Hu B.,
RA   Petrasch-Parwez E., Kilimann M.W.;
RT   "Neurobeachin: A protein kinase A-anchoring, beige/Chediak-Higashi
RT   protein homolog implicated in neuronal membrane traffic.";
RL   J. Neurosci. 20:8551-8565(2000).
DR   EMBL; Y18276; CAC18811.1; -.
SQ   SEQUENCE   2936 AA;  326738 MW;  A6CFD90CA666CEA4 CRC64;

  Query Match             1.5%;  Score 9;  DB 11;  Length 2936;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     357 VSGSRDATL 365
           |||||||||
Db    2730 VSGSRDATL 2738

RESULT  15
O66716
```

```
ID   O66716      PRELIMINARY;      PRT;   272 AA.
AC   O66716;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HYPOTHETICAL 31.6 KDA PROTEIN.
GN   AQ_397.
OS   Aquifex aeolicus.
OC   Bacteria; Aquificales; Aquificaceae; Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
DR   EMBL; AE000688; AAC06680.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   272 AA;  31561 MW;  652D06C62E15D581 CRC64;


  Query Match             1.4%;  Score 8;  DB 2;  Length 272;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    407 VWDPETET 414
          ||||||||
Db    161 VWDPETET 168


Search completed: September 25, 2001, 14:49:07
Job time: 211 sec
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          September 25, 2001, 14:38:25 ; Search time 20.43 Seconds
                                                (without alignments)
                                                1747.799 Million cell updates/sec

Title:           US-09-328-877A-8
Perfect score:   3143
Sequence:        1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:      412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
                1:  /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
               10:  /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
               11:  /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
               12:  /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
               13:  /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
               14:  /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
               15:  /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
               16:  /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
               17:  /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
               18:  /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
               19:  /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
               20:  /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
               21:  /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
               22:  /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                      %
Result             Query
  No.     Score    Match Length DB  ID                 Description
-----------------------------------------------------------------------------
    1      3143    100.0    589  20  AAY22466           Human mammary sel-
    2      3143    100.0    589  22  AAB59198           Human mammary sel-
    3      3140     99.9    589  21  AAB01204           Human GTPase assoc
    4      2983     94.9    559  20  AAY22467           Human mammary sel-
    5      2983     94.9    559  22  AAB59199           Human mammary sel-
    6    2890.5     92.0    627  20  AAY22461           Human hippocampal
    7    2890.5     92.0    627  22  AAB59193           Human hippocampal
    8    2890.5     92.0    666  20  AAY22471           Human C-term mychi
    9    2890.5     92.0    666  22  AAB59203           C-terminal mychis
   10    2890.5     92.0    669  20  AAY22470           Human Cterm V5 his
   11    2890.5     92.0    669  22  AAB59202           C-terminal V5 His
   12      2890     92.0    592  20  AAY22462           Human hippocampal
   13      2890     92.0    592  22  AAB59194           Human hippocampal
   14      2889     91.9    545  20  AAY22464           Human hippocampal
   15      2889     91.9    545  22  AAB59196           Human hippocampal
   16      2889     91.9    553  20  AAY22463           Human hippocampal
   17      2889     91.9    553  22  AAB59195           Human hippocampal
   18      2887     91.9    540  20  AAY22465           Human hippocampal
   19      2887     91.9    540  20  AAY22468           Human mammary sel-
   20      2887     91.9    540  22  AAB59197           Human hippocampal
   21      2887     91.9    540  22  AAB59200           Human mammary sel-
   22      2887     91.9    626  20  AAY22469           Human 6myc-N-sel-1
   23      2887     91.9    626  22  AAB59201           Protein encoded by
   24      1238     39.4    587  20  AAY03204           Amino acid sequenc
   25       650     20.7    640  21  AAY83252           F-box protein Met3
   26     641.5     20.4    542  21  AAY96696           Human E3 ubiquitin
   27       641     20.4    703  21  AAY83253           F-box protein Pop2
   28       640     20.4    569  20  AAY24054           A human beta-trans
   29       640     20.4    569  21  AAB12813           Human beta-transdu
   30       640     20.4    569  21  AAY96697           Human beta-TrCP.
   31       640     20.4    569  21  AAY83041           F-box protein FBP-
   32       640     20.4    569  21  AAY83250           F-box protein hBet
   33       640     20.4    569  21  AAY44249           Human cell signall
   34       637     20.3    569  22  AAB48298           Human ZF11 protein
   35       635     20.2    569  21  AAB12812           Mouse ubiquitin li
   36       635     20.2    569  21  AAY83254           F-box protein FWD1
   37       623     19.8    517  16  AAR85852           WD-40 domain-contg
   38     618.5     19.7    779  16  AAR85854           WD-40 domain-contg
   39     618.5     19.7    779  21  AAY83249           F-box protein Cdc4
   40       461     14.7    317  21  AAB63186           Human secreted pro
   41       461     14.7    334  22  AAB68529           Human GTP-binding
   42       439     14.0    409  16  AAR70002           OPDE 45 kDa subuni
   43       439     14.0    410  16  AAR70005           OPDE 45 kDa subuni
   44     424.5     13.5    409  16  AAR85868           WD-40 domain-contg
   45       396     12.6    323  21  AAB63185           Gene 3 human secre

                              ALIGNMENTS

RESULT   1
AAY22466
ID   AAY22466 standard; Protein; 589 AA.
XX
AC   AAY22466;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human mammary sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;
KW   mammary gland; therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99702.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 60-63; 91pp; English.
XX
```

```
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human mammary sel-10 protein. The polypeptides
CC   can be used to alter presenilin function. Compounds which inhibit either
CC   the expression or the activity of the human sel-10 polypeptides may
CC   reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and
CC   therefore may be useful for the prevention or treatment of Alzheimer's
CC   disease.
XX
SQ   Sequence   589 AA;

  Query Match             100.0%;  Score 3143;  DB 20;  Length 589;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-298;
  Matches  589;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MSKPGKPTLNHGLVPVDLKSAKEPLPHQTVMKIFSISIIAQGLPFCRRRMKRKLDHGSEV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 mskpgkptlnhglvpvdlksakeplphqtvmkifsisiiaqglpfcrrrmkrkldhgsev 60

Qy       61 RSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 rsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrritsvqpptglqewl 120

Qy      121 KMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 kmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpkelalyvlsflep 180

Qy      181 KDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFIHSPWKSAYIRQ 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 kdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfihspwksayirq 240

Qy      241 HRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 hridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvwsavtgkclrtlv 300

Qy      301 GHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcmhlhekrvvsgs 360

Qy      361 RDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 rdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwdpetetclhtlq 420

Qy      421 GHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 ghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgmelkdnilvsgn 480

Qy      481 ADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIR 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 adstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvklwdlktgefir 540

Qy      541 NLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 nlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 589


RESULT   2
AAB59198
ID   AAB59198 standard; protein; 589 AA.
XX
AC   AAB59198;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human mammary sel-10-1 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 79-82; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   589 AA;

  Query Match             100.0%;  Score 3143;  DB 22;  Length 589;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-298;
  Matches  589;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MSKPGKPTLNHGLVPVDLKSAKEPLPHQTVMKIFSISIIAQGLPFCRRRMKRKLDHGSEV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 mskpgkptlnhglvpvdlksakeplphqtvmkifsisiiaqglpfcrrrmkrkldhgsev 60

Qy       61 RSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 rsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrritsvqpptglqewl 120

Qy      121 KMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEP 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 kmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpkelalyvlsflep 180

Qy      181 KDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFIHSPWKSAYIRQ 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 kdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfihspwksayirq 240

Qy      241 HRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 hridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvwsavtgkclrtlv 300

Qy      301 GHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcmhlhekrvvsgs 360

Qy      361 RDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 rdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwdpetetclhtlq 420

Qy      421 GHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 ghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgmelkdnilvsgn 480

Qy      481 ADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIR 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 adstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvklwdlktgefir 540

Qy      541 NLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
              |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    541 nlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 589

RESULT   3
AAB01204
ID   AAB01204 standard; Protein; 589 AA.
XX
AC   AAB01204;
XX
DT   03-NOV-2000  (first entry)
XX
DE   Human GTPase associated protein-29.
XX
KW   Guanine nucleotide binding protein; GTP-binding protein; G-protein;
KW   GTPase; GTPase associated protein; GTPAP; cell proliferation;
KW   autoimmune; inflammatory; immune system disorder; cancer; AIDS;
KW   acquired immune deficiency syndrome; asthma; atherosclerosis;
KW   arthritis; systemic lupus erythematosus; psoriasis; human.
XX
OS   Homo sapiens.
XX
PN   WO200031263-A2.
XX
PD   02-JUN-2000.
XX
PF   23-NOV-1999;   99WO-US28013.
XX
PR   23-NOV-1998;   98US-0109592.
PR   04-FEB-1999;   99US-0118610.
PR   06-APR-1999;   99US-0127990.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL,  Tang YT,  Bandman O,  Lal P,  Yue H,  Lu DAM,  Baughn MR;
PI   Yang J,  Azimzai Y;
XX
DR   WPI; 2000-400073/34.
DR   N-PSDB; AAA49199.
XX
PT   Human GTPase associated proteins, polynucleotides, and antibodies,
PT   useful for diagnosing, preventing and treating various diseases such as
PT   atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),
PT   asthma, and autoimmune diseases -
XX
PS   Claim 1; Page 118-120; 144pp; English.
XX
CC   Human cDNA libraries from various tissues were screened for GTPase
CC   associated proteins (GTPAP). The present sequence is human
CC   GTPAP-29 protein. This sequence was derived from a cDNA library of
CC   brain tumour tissue. This protein is expressed in nervous,
CC   reproductive and gastrointestinal tissue. The GTPAP proteins may be
CC   used to define agonists and antagonists of GTPAP activity and to
CC   generate antibodies to GTPAP. This means the GTPAP proteins may be
CC   useful for treatment or prevention of diseases associated with GTPAP
CC   such as cell proliferation disorders, autoimmune disorders,
CC   inflammatory disorders, immune system disorders, cancer, AIDS, asthma,
CC   atherosclerosis, arthritis, systemic lupus erythematosus and psoriasis.
XX
SQ   Sequence   589 AA;

  Query Match             99.9%;  Score 3140;  DB 21;  Length 589;
  Best Local Similarity   99.8%;  Pred. No. 2.2e-298;
  Matches  588;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MSKPGKPTLNHGLVPVDLKSAKEPLPHQTVMKIFSISIIAQGLPFCRRRMKRKLDHGSEV 60
          |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db      1 mskpgkptlnhglvpvdlksakeplphqtvmrifsisiiaqglpfcrrrmkrkldhgsev 60

Qy     61 RSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 rsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrritsvqpptglqewl 120

Qy    121 KMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEP 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 kmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpkelalyvlsflep 180

Qy    181 KDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFIHSPWKSAYIRQ 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 kdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfihspwksayirq 240

Qy    241 HRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLV 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 hridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvwsavtgkclrtlv 300

Qy    301 GHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcmhlhekrvvsgs 360

Qy    361 RDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 rdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwdpetetclhtlq 420

Qy    421 GHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGN 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgmelkdnilvsgn 480

Qy    481 ADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIR 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 adstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvklwdlktgefir 540

Qy    541 NLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 nlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 589

RESULT   4
AAY22467
ID   AAY22467 standard; Protein; 559 AA.
XX
AC   AAY22467;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human mammary sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;
KW   mammary gland; therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99702.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 63-66; 91pp; English.
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human mammary sel-10 protein. The polypeptides
CC   can be used to alter presenilin function. Compounds which inhibit either
CC   the expression or the activity of the human sel-10 polypeptides may
CC   reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and
CC   therefore may be useful for the prevention or treatment of Alzheimer's
```

```
CC   disease.
XX
SQ   Sequence   559 AA;

  Query Match             94.9%;  Score 2983;  DB 20;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-283;
  Matches  559;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       31 MKIFSISIIAQGLPFCRRRMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTT 90
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 mkifsisiiaqglpfcrrrmkrkldhgsevrsfslgkkpckvseytsttglvpcsatptt 60

Qy       91 FGDLRAANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHM 150
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 fgdlraangqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhm 120

Qy      151 MQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEE 210
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 mqviepqfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckee 180

Qy      211 GIDEPLHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQ 270
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 gideplhikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclq 240

Qy      271 FCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNA 330
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 fcgnrivsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwna 300

Qy      331 ETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQ 390
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 etgecihtlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvq 360

Qy      391 YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDV 450
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 ydgrrvvsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdv 420

Qy      451 ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVT 510
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 etgncihtltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavt 480

Qy      511 CLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGS 570
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 clqfnknfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgs 540

Qy      571 RNGTEETKLLVLDFDVDMK 589
            |||||||||||||||||||
Db      541 rngteetkllvldfdvdmk 559

RESULT   5
AAB59199
ID   AAB59199 standard; protein; 559 AA.
XX
AC   AAB59199;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human mammary sel-10-2 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;    99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 83-86; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   559 AA;

  Query Match             94.9%;  Score 2983;  DB 22;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-283;
  Matches  559;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       31 MKIFSISIIAQGLPFCRRRMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTT 90
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 mkifsisiiaqglpfcrrrmkrkldhgsevrsfslgkkpckvseytsttglvpcsatptt 60

Qy       91 FGDLRAANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHM 150
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 fgdlraangqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhm 120

Qy      151 MQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEE 210
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 mqviepqfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckee 180

Qy      211 GIDEPLHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQ 270
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 gideplhikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclq 240

Qy      271 FCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNA 330
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 fcgnrivsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwna 300

Qy      331 ETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQ 390
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 etgecihtlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvq 360

Qy      391 YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDV 450
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 ydgrrvvsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdv 420

Qy      451 ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVT 510
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 etgncihtltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavt 480

Qy      511 CLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGS 570
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 clqfnknfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgs 540

Qy      571 RNGTEETKLLVLDFDVDMK 589
            |||||||||||||||||||
Db      541 rngteetkllvldfdvdmk 559

RESULT   6
AAY22461
ID   AAY22461 standard; Protein; 627 AA.
```

```
XX
AC   AAY22461;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human hippocampal sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99701.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 43-46; 91pp; English.
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human hippocampal sel-10 protein. The
CC   polypeptides can be used to alter presenilin function. Compounds which
CC   inhibit either the expression or the activity of the human sel-10
CC   polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)
CC   or PS-2, and therefore may be useful for the prevention or treatment of
CC   Alzheimer's disease.
XX
SQ   Sequence   627 AA;

  Query Match            92.0%;  Score 2890.5;  DB 20;  Length 627;
  Best Local Similarity  89.9%;  Pred. No. 6.3e-274;
  Matches  551;  Conservative   11;  Mismatches   10;  Indels   41;  Gaps    4;

Qy     13 LVPVDLKSAKEPLPHQTVMKIFSISIIA----QGLPFCRR-------------------- 48
          |:|| | :    ||  | : : ::  :     :|| :|:|
Db     20 llpvllpn----lpfltclsmstlesvtylpekgl-ycqrlpssrthggteslkgknten 74

Qy     49 ------------RMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRA 96
                      :|||||||||||||||||||||||||||||||||||||||||||||
Db     75 mgfygtlkmifykmkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlra 134

Qy     97 ANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEP 156
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 angqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviep 194

Qy    157 QFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPL 216
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 qfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckeegidepl 254

Qy    217 HIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRI 276
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 hikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnri 314

Qy    277 VSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECI 336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 vsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgeci 374

Qy    337 HTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRV 396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 htlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrv 434
```

```
Qy    397 VSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCI 456
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 vsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgnci 494

Qy    457 HTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNK 516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 htltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnk 554

Qy    517 NFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEE 576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 nfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngtee 614

Qy    577 TKLLVLDFDVDMK 589
          |||||||||||||
Db    615 tkllvldfdvdmk 627

RESULT   7
AAB59193
ID   AAB59193 standard; protein; 627 AA.
XX
AC   AAB59193;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human hippocampal sel-10-1 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 60-64; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   627 AA;

  Query Match            92.0%;  Score 2890.5;  DB 22;  Length 627;
  Best Local Similarity  89.9%;  Pred. No. 6.3e-274;
  Matches  551;  Conservative   11;  Mismatches   10;  Indels   41;  Gaps    4;

Qy     13 LVPVDLKSAKEPLPHQTVMKIFSISIIA----QGLPFCRR-------------------- 48
          |:|| | :    ||  | : : ::  :     :|| :|:|
Db     20 llpvllpn----lpfltclsmstlesvtylpekgl-ycqrlpssrthggteslkgknten 74
```

```
Qy         49 ------------RMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRA 96
                          :|||||||||||||||||||||||||||||||||||||||||||||||
Db         75 mgfygtlkmifykmkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlra 134

Qy         97 ANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEP 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 angqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviep 194

Qy        157 QFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPL 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 qfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckeegidepl 254

Qy        217 HIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRI 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 hikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnri 314

Qy        277 VSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECI 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 vsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgeci 374

Qy        337 HTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRV 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 htlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrv 434

Qy        397 VSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCI 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 vsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgnci 494

Qy        457 HTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNK 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 htltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnk 554

Qy        517 NFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEE 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 nfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngtee 614

Qy        577 TKLLVLDFDVDMK 589
              |||||||||||||
Db        615 tkllvldfdvdmk 627

RESULT   8
AAY22471
ID   AAY22471 standard; Protein; 666 AA.
XX
AC   AAY22471;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human C-term mychis tagged sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW   therapy; mammary gland.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99716.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Example 3; Page 85-88; 91pp; English.
XX
CC   This sequence is an epitope-tagged version of a human sel-10 protein
CC   of the invention. The human sel-10 proteins of the invention are isolated
CC   from hippocampus and mammary gland. The polypeptides can be used to
CC   alter presenilin function. Compounds which inhibit either the expression
CC   or the activity of the human sel-10 polypeptides may reverse the effects
CC   of mutations to presenilin-1 (PS-1) or PS-2, and therefore may be useful
CC   for the prevention or treatment of Alzheimer's disease.
XX
SQ   Sequence   666 AA;


  Query Match             92.0%;  Score 2890.5;  DB 20;  Length 666;
  Best Local Similarity   89.9%;  Pred. No. 7e-274;
  Matches  551;  Conservative   11;  Mismatches   10;  Indels   41;  Gaps    4;

Qy         13 LVPVDLKSAKEPLPHQTVMKIFSISIIA----QGLPFCRR-------------------- 48
              |:|| | :    ||  | : : ::  :     :|| :|:|
Db         20 llpvllpn----lpfltclsmstlesvtylpekgl-ycqrlpssrthggteslkgknten 74

Qy         49 ------------RMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRA 96
                          :|||||||||||||||||||||||||||||||||||||||||||||||
Db         75 mgfygtlkmifykmkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlra 134

Qy         97 ANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEP 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 angqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviep 194

Qy        157 QFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPL 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 qfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckeegidepl 254

Qy        217 HIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRI 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 hikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnri 314

Qy        277 VSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECI 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 vsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgeci 374

Qy        337 HTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRV 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 htlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrv 434

Qy        397 VSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCI 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 vsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgnci 494

Qy        457 HTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNK 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 htltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnk 554

Qy        517 NFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEE 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 nfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngtee 614

Qy        577 TKLLVLDFDVDMK 589
              |||||||||||||
Db        615 tkllvldfdvdmk 627

RESULT   9
AAB59203
ID   AAB59203 standard; protein; 666 AA.
XX
AC   AAB59203;
XX
DT   23-MAR-2001  (first entry)
XX
DE   C-terminal mychis tagged sel-10 protein.
```

```
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Example 3; Page 108-112; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   666 AA;

  Query Match             92.0%;  Score 2890.5;  DB 22;  Length 666;
  Best Local Similarity   89.9%;  Pred. No. 7e-274;
  Matches  551;  Conservative   11;  Mismatches   10;  Indels   41;  Gaps    4;

Qy      13 LVPVDLKSAKEPLPHQTVMKIFSISIIA----QGLPFCRR-------------------- 48
           |:|| | :    ||  | : : ::  :     :|| :|:|
Db      20 llpvllpn----lpfltclsmstlesvtylpekgl-ycqrlpssrthggteslkgknten 74

Qy      49 ------------RMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRA 96
                       :|||||||||||||||||||||||||||||||||||||||||||||||
Db      75 mgfygtlkmifykmkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlra 134

Qy      97 ANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEP 156
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 angqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviep 194

Qy     157 QFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPL 216
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 qfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckeegidepl 254

Qy     217 HIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRI 276
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 hikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnri 314

Qy     277 VSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECI 336
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 vsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgeci 374

Qy     337 HTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRV 396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 htlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrv 434

Qy     397 VSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCI 456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     435 vsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgnci 494

Qy     457 HTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNK 516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     495 htltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnk 554

Qy     517 NFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEE 576
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     555 nfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngtee 614

Qy     577 TKLLVLDFDVDMK 589
           |||||||||||||
Db     615 tkllvldfdvdmk 627

RESULT  10
AAY22470
ID   AAY22470 standard; Protein; 669 AA.
XX
AC   AAY22470;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human Cterm V5 his tagged sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW   therapy; mammary gland.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99715.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Example 3; Page 79-83; 91pp; English.
XX
CC   This sequence is an epitope-tagged version of a human sel-10 protein
CC   of the invention. The human sel-10 proteins of the invention are isolated
CC   from hippocampus and mammary gland. The polypeptides can be used to
CC   alter presenilin function. Compounds which inhibit either the expression
CC   or the activity of the human sel-10 polypeptides may reverse the effects
CC   of mutations to presenilin-1 (PS-1) or PS-2, and therefore may be useful
CC   for the prevention or treatment of Alzheimer's disease.
XX
SQ   Sequence   669 AA;

  Query Match             92.0%;  Score 2890.5;  DB 20;  Length 669;
  Best Local Similarity   89.9%;  Pred. No. 7e-274;
  Matches  551;  Conservative   11;  Mismatches   10;  Indels   41;  Gaps    4;

Qy      13 LVPVDLKSAKEPLPHQTVMKIFSISIIA----QGLPFCRR-------------------- 48
           |:|| | :    ||  | : : ::  :     :|| :|:|
Db      20 llpvllpn----lpfltclsmstlesvtylpekgl-ycqrlpssrthggteslkgknten 74

Qy      49 ------------RMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRA 96
                       :|||||||||||||||||||||||||||||||||||||||||||||||
Db      75 mgfygtlkmifykmkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlra 134
```

```
Qy       97 ANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEP 156
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      135 angqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviep 194

Qy      157 QFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPL 216
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      195 qfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckeegidepl 254

Qy      217 HIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRI 276
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      255 hikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnri 314

Qy      277 VSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECI 336
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      315 vsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgeci 374

Qy      337 HTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRV 396
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      375 htlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrv 434

Qy      397 VSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCI 456
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      435 vsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgnci 494

Qy      457 HTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNK 516
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      495 htltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnk 554

Qy      517 NFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEE 576
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      555 nfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngtee 614

Qy      577 TKLLVLDFDVDMK 589
            |||||||||||||
Db      615 tkllvldfdvdmk 627
```

```
RESULT 11
AAB59202
ID   AAB59202 standard; protein; 669 AA.
XX
AC   AAB59202;
XX
DT   23-MAR-2001  (first entry)
XX
DE   C-terminal V5 His tagged sel-10 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Example 3; Page 101-105; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   669 AA;

  Query Match             92.0%;  Score 2890.5;  DB 22;  Length 669;
  Best Local Similarity   89.9%;  Pred. No. 7e-274;
  Matches  551;  Conservative   11;  Mismatches   10;  Indels   41;  Gaps    4;

Qy       13 LVPVDLKSAKEPLPHQTVMKIFSISIIA----QGLPFCRR-------------------- 48
            |:|| | :    ||  | : : ::  :     :|| :|:|
Db       20 llpvllpn----lpfltclsmstlesvtylpekgl-ycqrlpssrthggteslkgknten 74

Qy       49 ------------RMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRA 96
                        :|||||||||||||||||||||||||||||||||||||||||||||||
Db       75 mgfygtlkmifykmkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlra 134

Qy       97 ANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEP 156
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      135 angqgqqrrritsvqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviep 194

Qy      157 QFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPL 216
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      195 qfqrdfisllpkelalyvlsflepkdllqaaqtcrywrilaednllwrekckeegidepl 254

Qy      217 HIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRI 276
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      255 hikrrkvikpgfihspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnri 314

Qy      277 VSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECI 336
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      315 vsgsddntlkvwsavtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgeci 374

Qy      337 HTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRV 396
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      375 htlyghtstvrcmhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrv 434

Qy      397 VSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCI 456
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      435 vsgaydfmvkvwdpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgnci 494

Qy      457 HTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNK 516
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      495 htltghqsltsgmelkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnk 554

Qy      517 NFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEE 576
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      555 nfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngtee 614

Qy      577 TKLLVLDFDVDMK 589
            |||||||||||||
Db      615 tkllvldfdvdmk 627
```

```
RESULT 12
AAY22462
ID   AAY22462 standard; Protein; 592 AA.
XX
AC   AAY22462;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human hippocampal sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW   therapy.
```

```
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99701.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 47-50; 91pp; English.
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human hippocampal sel-10 protein. The
CC   polypeptides can be used to alter presenilin function. Compounds which
CC   inhibit either the expression or the activity of the human sel-10
CC   polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)
CC   or PS-2, and therefore may be useful for the prevention or treatment of
CC   Alzheimer's disease.
XX
SQ   Sequence   592 AA;

  Query Match             92.0%;  Score 2890;  DB 20;  Length 592;
  Best Local Similarity   93.4%;  Pred. No. 6.5e-274;
  Matches  549;  Conservative    3;  Mismatches   10;  Indels   26;  Gaps    2;

Qy       2 SKPGKPTLNHGLVPVDLKSAKEPLPHQTVMKIFSISIIAQGLPFCRRRMKRKLDHGSEVR 61
           |  || | | |              : |:  ||             :|||||||||||||
Db      31 slkgkntenmgf-------------ygtlkmif-------------ykmkrkldhgsevr 64

Qy      62 SFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWLK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 sfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrritsvqpptglqewlk 124

Qy     122 MFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPK 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125 mfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpkelalyvlsflepk 184

Qy     182 DLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFIHSPWKSAYIRQH 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185 dllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfihspwksayirqh 244

Qy     242 RIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVG 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     245 ridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvwsavtgkclrtlvg 304

Qy     302 HTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSR 361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     305 htggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcmhlhekrvvsgsr 364

Qy     362 DATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQG 421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     365 datlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwdpetetclhtlqg 424

Qy     422 HTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNA 481
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     425 htnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgmelkdnilvsgna 484

Qy     482 DSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIRN 541
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     485 dstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvklwdlktgefirn 544
```

```
Qy     542 LVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     545 lvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 592

RESULT 13
AAB59194
ID   AAB59194 standard; protein; 592 AA.
XX
AC   AAB59194;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human hippocampal sel-10-2 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
PT   -
XX
PS   Claim 1; Page 64-68; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   592 AA;

  Query Match             92.0%;  Score 2890;  DB 22;  Length 592;
  Best Local Similarity   93.4%;  Pred. No. 6.5e-274;
  Matches  549;  Conservative    3;  Mismatches   10;  Indels   26;  Gaps    2;

Qy       2 SKPGKPTLNHGLVPVDLKSAKEPLPHQTVMKIFSISIIAQGLPFCRRRMKRKLDHGSEVR 61
           |  || | | |              : |:  ||             :|||||||||||||
Db      31 slkgkntenmgf-------------ygtlkmif-------------ykmkrkldhgsevr 64

Qy      62 SFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWLK 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 sfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrritsvqpptglqewlk 124

Qy     122 MFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPK 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125 mfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpkelalyvlsflepk 184

Qy     182 DLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFIHSPWKSAYIRQH 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185 dllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgfihspwksayirqh 244
```

```
Qy     242 RIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVG 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     245 ridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvwsavtgkclrtlvg 304

Qy     302 HTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSR 361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     305 htggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrcmhlhekrvvsgsr 364

Qy     362 DATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQG 421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     365 datlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvwdpetetclhtlqg 424

Qy     422 HTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNA 481
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     425 htnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgmelkdnilvsgna 484

Qy     482 DSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIRN 541
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     485 dstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtvklwdlktgefirn 544

Qy     542 LVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK 589
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     545 lvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 592


RESULT  14
AAY22464
ID   AAY22464 standard; Protein; 545 AA.
XX
AC   AAY22464;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human hippocampal sel-10 protein sequence.
XX
KW   Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9932623-A1.
XX
PD   01-JUL-1999.
XX
PF   17-DEC-1998;   98WO-US26820.
XX
PR   19-DEC-1997;   97US-0068243.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME,  Li J,  Pauley AM;
XX
DR   WPI; 1999-458026/38.
DR   N-PSDB; AAX99701.
XX
PT   New isolated human sel-10 polypeptides
XX
PS   Claim 24; Page 53-56; 91pp; English.
XX
CC   This sequence represents a human sel-10 protein of the invention. This
CC   sequence is specifically a human hippocampal sel-10 protein. The
CC   polypeptides can be used to alter presenilin function. Compounds which
CC   inhibit either the expression or the activity of the human sel-10
CC   polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)
CC   or PS-2, and therefore may be useful for the prevention or treatment of
CC   Alzheimer's disease.
XX
SQ   Sequence   545 AA;


  Query Match             91.9%;  Score 2889;  DB 20;  Length 545;
  Best Local Similarity   99.8%;  Pred. No. 7.1e-274;
  Matches  540;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      49 RMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRIT 108
           :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5 kmkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrit 64

Qy     109 SVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPK 168
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 svqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpk 124

Qy     169 ELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGF 228
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125 elalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgf 184

Qy     229 IHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVW 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185 ihspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvw 244

Qy     289 SAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRC 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     245 savtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrc 304

Qy     349 MHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVW 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     305 mhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvw 364

Qy     409 DPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSG 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     365 dpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsg 424

Qy     469 MELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTV 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     425 melkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtv 484

Qy     529 KLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDM 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     485 klwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdm 544

Qy     589 K 589
           |
Db     545 k 545


RESULT  15
AAB59196
ID   AAB59196 standard; protein; 545 AA.
XX
AC   AAB59196;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Human hippocampal sel-10-4 protein.
XX
KW   Sel-10; human; Alzheimer's disease; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200075328-A1.
XX
PD   14-DEC-2000.
XX
PF   23-MAY-2000; 2000WO-US09814.
XX
PR   09-JUN-1999;   99US-0328877.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Pauley AM,  Gurney ME,  Li J;
XX
DR   WPI; 2001-102404/11.
XX
PT   New human sel-10 polypeptides and their encoding polynucleotides,
PT   useful for raising antibodies for detecting sel-10 polypeptide
PT   expression and as drug targets in the treatment of Alzheimer's disease
```

```
PT   -
XX
PS   Claim 1; Page 72-75; 116pp; English.
XX
CC   The present invention relates to human sel-10. The sel-10 proteins of
CC   the invention are useful for raising monoclonal or polyclonal
CC   antibodies useful in diagnostic assays for detecting sel-10
CC   polypeptide expression. The sel-10 polypeptides are also useful as drug
CC   targets for decreasing antibody levels in the treatment of Alzheimer's
CC   disease. It is also useful for identifying agents capable of
CC   altering the production level of Abeta. The polynucleotides are useful
CC   for developing assays for identifying agents capable of interfering
CC   with the biological pathways that lead to Alzheimer's disease.
XX
SQ   Sequence   545 AA;

  Query Match             91.9%;  Score 2889;  DB 22;  Length 545;
  Best Local Similarity   99.8%;  Pred. No. 7.1e-274;
  Matches  540;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 RMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRIT 108
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 kmkrkldhgsevrsfslgkkpckvseytsttglvpcsatpttfgdlraangqgqqrrrit 64

Qy        109 SVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPK 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 svqpptglqewlkmfqswsgpekllaldelidsceptqvkhmmqviepqfqrdfisllpk 124

Qy        169 ELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGF 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 elalyvlsflepkdllqaaqtcrywrilaednllwrekckeegideplhikrrkvikpgf 184

Qy        229 IHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVW 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ihspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrivsgsddntlkvw 244

Qy        289 SAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRC 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 savtgkclrtlvghtggvwssqmrdniiisgstdrtlkvwnaetgecihtlyghtstvrc 304

Qy        349 MHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVW 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 mhlhekrvvsgsrdatlrvwdietgqclhvlmghvaavrcvqydgrrvvsgaydfmvkvw 364

Qy        409 DPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 dpetetclhtlqghtnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsg 424

Qy        469 MELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTV 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 melkdnilvsgnadstvkiwdiktgqclqtlqgpnkhqsavtclqfnknfvitssddgtv 484

Qy        529 KLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDM 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 klwdlktgefirnlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdm 544

Qy        589 K 589
              |
Db        545 k 545


Search completed: September 25, 2001, 14:41:10
Job time: 165 sec
```

THIS PAGE BLANK (USPTO)

```
                     GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          September 25, 2001, 14:39:00 ; Search time 13.65 Seconds
                                           (without alignments)
                                           888.477 Million cell updates/sec

Title:           US-09-328-877A-8
Perfect score:   3143
Sequence:        1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
                   %
Result           Query
  No.   Score    Match Length DB  ID                          Description
 -----------------------------------------------------------------------------
    1    1238    39.4    587  3  US-08-899-578-2             Sequence 2, Appli
    2     671    21.3    209  3  US-08-899-578-6             Sequence 6, Appli
    3     623    19.8    517  1  US-08-190-802A-30           Sequence 30, Appl
    4   618.5    19.7    779  1  US-08-190-802A-32           Sequence 32, Appl
    5     439    14.0    409  2  US-08-283-917-3             Sequence 3, Appli
    6     439    14.0    409  2  US-08-961-716-3             Sequence 3, Appli
    7     439    14.0    410  2  US-08-283-917-9             Sequence 9, Appli
    8     439    14.0    410  2  US-08-961-716-9             Sequence 9, Appli
    9   424.5    13.5    409  1  US-08-190-802A-51           Sequence 51, Appl
   10     410    13.0    212  3  US-08-899-578-7             Sequence 7, Appli
   11   376.5    12.0    422  1  US-08-190-802A-52           Sequence 52, Appl
   12   364.5    11.6    514  1  US-08-190-802A-66           Sequence 66, Appl
   13     332    10.6    704  1  US-08-188-582-5             Sequence 5, Appli
   14     332    10.6    704  1  US-08-646-715-5             Sequence 5, Appli
   15     326    10.4    704  1  US-08-190-802A-62           Sequence 62, Appl
   16     326    10.4    704  2  US-08-308-818-3             Sequence 3, Appli
   17   319.5    10.2    713  1  US-08-190-802A-63           Sequence 63, Appl
   18   317.5    10.1    704  1  US-08-188-582-18            Sequence 18, Appl
   19   317.5    10.1    704  1  US-08-646-715-18            Sequence 18, Appl
   20   311.5     9.9    375  4  US-09-063-743-1             Sequence 1, Appli
   21     306     9.7    798  1  US-08-190-802A-64           Sequence 64, Appl
   22     306     9.7    798  1  US-08-190-802A-68           Sequence 68, Appl
   23     306     9.7    798  2  US-08-308-818-2             Sequence 2, Appli
   24     288     9.2    318  1  US-08-190-802A-33           Sequence 33, Appl
   25     287     9.1    906  1  US-08-190-802A-31           Sequence 31, Appl
   26   280.5     8.9    340  1  US-08-190-802A-38           Sequence 38, Appl
   27     276     8.8    340  4  US-09-180-783-2             Sequence 2, Appli
   28   272.5     8.7    340  1  US-08-190-802A-42           Sequence 42, Appl
   29   269.5     8.6    325  1  US-07-626-589-2             Sequence 2, Appli
   30   269.5     8.6    325  1  US-08-190-802A-57           Sequence 57, Appl
   31   269.5     8.6    325  1  US-08-323-444A-1            Sequence 1, Appli
   32   269.5     8.6    325  1  US-08-236-410-2             Sequence 2, Appli
   33   269.5     8.6    325  1  US-08-465-421-2             Sequence 2, Appli
   34     269     8.6    343  4  US-09-063-743-5             Sequence 5, Appli
   35     266     8.5    305  3  US-08-965-600-1             Sequence 1, Appli
   36   265.5     8.4    326  1  US-08-190-802A-39           Sequence 39, Appl
   37   259.5     8.3    395  3  US-09-032-372-1             Sequence 1, Appli
   38     256     8.1   2627  2  US-08-751-189-3             Sequence 3, Appli
   39     256     8.1   2627  2  US-09-060-836-3             Sequence 3, Appli
   40     256     8.1   2627  4  US-09-184-445-3             Sequence 3, Appli
   41     253     8.0    340  1  US-08-190-802A-40           Sequence 40, Appl
   42     250     8.0    341  1  US-08-190-802A-45           Sequence 45, Appl
   43   246.5     7.8    317  1  US-08-190-802A-27           Sequence 27, Appl
   44   246.5     7.8    317  1  US-08-190-802A-41           Sequence 41, Appl
   45   246.5     7.8    317  1  US-08-190-802A-47           Sequence 47, Appl

                               ALIGNMENTS

RESULT   1
US-08-899-578-2
; Sequence 2, Application US/08899578
; Patent No. 6087153
;  GENERAL INFORMATION:
;    APPLICANT:  Greenwald, Iva
;    APPLICANT:  Hubbard, E. Jane
;    TITLE OF INVENTION:  SEL-10 AND USES THEREOF
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/899,578
;      FILING DATE:  24-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/53200/JPW/AKC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 278-0525
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 587 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-899-578-2


  Query Match             39.4%;  Score 1238;  DB 3;  Length 587;
  Best Local Similarity   45.6%;  Pred. No. 2:1e-116;
  Matches  262;  Conservative   85;  Mismatches  162;  Indels   66;  Gaps   11;

Qy      55 DHGSEVRSFSLGKKPC----KVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSV 110
           |:| |  |:| |        |:|        :  ||:|:   ||
Db      31 DNGEE-SSYSNGSSSSYNADKLSSSRPLQHKLDLSASPSRNNDLNPR------------- 76

Qy     111 QPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKEL 170
```

```
               ::  : :|:  |  |:: |   |:     | :: :  :||| |||||:| || ||
Db     77 -----VEHLIALFKDLSSAEQMDAFTRLLQESNMTNIRQLRAIIEPHFQRDFLSCLPVEL 131

Qy    171 ALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEE-------------------- 210
           : :|  |   |||: ||  : |::::| : :|:    ||
Db    132 GMKILHNLTGYDLLKVAQVSKNWKLISEIDKIWKSLGVEEFKHHPDPTDRVTGAWQGTAI 191

Qy    211 --GIDEPLHI--------KRRKVIKPG--FIHSPWKSAYIRQHRIDTNWRRGELKSPKVL 258
            |:  | ||        :  |: | |  |  : ||  |:|  :|: ||    :    ||
Db    192 AAGVTIPDHIQPCDLNVHRFLKLQKFGDIFERAADKSRYLRADKIEKNWNANPIMGSAVL 251

Qy    259 KGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRD--NII 316
          :||:||||||:|   : :|:||||||||||    |: : ||||||||||:||:      |
Db    252 RGHEDHVITCMQIHDDVLVTGSDDNTLKVWCIDKGEVMYTLVGHTGGVWTSQISQCGRYI 311

Qy    317 ISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL 376
          :|||||||:|||:   |  :||| |||||||||  :    :|:|||| ||||||:|:|: |
Db    312 VSGSTDRTVKVWSTVDGSLLHTLQGHTSTVRCMAMAGSILVTGSRDTTLRVWDVESGRHL 371

Qy    377 HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIH- 435
            | || |||||||:||  |||| ||| ||:|:  |  |: || || |||||| |:
Db    372 ATLHGHHAAVRCVQFDGTTVVSGGYDFTVKIWNAHTGRCIRTLTGHNNRVYSLLFESERS 431

Qy    436 -VVSGSLDTSIRVWDV---ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIK 491
           | ||||||||||||    |   |:  | ||  ||||||:|: ||||| |||| |::|||
Db    432 IVCSGSLDTSIRVWDFTRPEGQECVALLQGHTSLTSGMQLRGNILVSCNADSHVRVWDIH 491

Qy    492 TGQCLQTLQGPNKHQSAVTCLQ-FNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGS 550
           | |:  | |   |:||:| || | :| | ||||||||||||:: |  ||:||||:|||:
Db    492 EGTCVHMLSG---HRSAITSLQWFGRNMVATSSDDGTVKLWDIERGALIRDLVTLDSGGN 548

Qy    551 GGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFD 585
          || :||: :::| | ||||||| |||||:::||||
Db    549 GGCIWRLCSTSTMLACAVGSRNNTEETKVILLDFD 583


RESULT   2
US-08-899-578-6
; Sequence 6, Application US/08899578
; Patent No. 6087153
;  GENERAL INFORMATION:
;    APPLICANT:  Greenwald, Iva
;    APPLICANT:  Hubbard, E. Jane
;    TITLE OF INVENTION:  SEL-10 AND USES THEREOF
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/899,578
;      FILING DATE:  24-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/53200/JPW/AKC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 278-0525
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  209 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-899-578-6

  Query Match             21.3%;  Score 671;  DB 3;  Length 209;
  Best Local Similarity   52.2%;  Pred. No. 8.3e-60;
  Matches  145;  Conservative   24;  Mismatches   35;  Indels   74;  Gaps    9;

Qy    260 GHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRD--NIII 317
          ||:||||||:|   : :|:||||||||||            |||||||:||:      |:
Db      1 GHEDHVITCMQIHDDVLVTGSDDNTLKVW-----------CGHTGGVWTSQISQCGRYIV 49

Qy    318 SGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLH 377
          |||||||:|||:           |||||||||  :    :|:|||| ||||||
Db     50 SGSTDRTVKVWS-----------GHTSTVRCMAMAGSILVTGSRDTTLRVWD-------- 90

Qy    378 VLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIH-- 435
             || |||||||:||  |||| ||| ||:|:           || |||||| |:
Db     91 ---GHHAAVRCVQFDGTTVVSGGYDFTVKIWN-----------GHNNRVYSLLFESERSI 136

Qy    436 VVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQC 495
          | ||||||||||||           || ||||||:|: ||||| |||| |::||
Db    137 VCSGSLDTSIRVWD-----------GHTSLTSGMQLRGNILVSCNADSHVRVWD------ 179

Qy    496 LQTLQGPNKHQSAVTCLQ-FNKNFVITSSDDGTVKLWD 532
                   |:||:| || | :| | ||||||||||||
Db    180 --------GHRSAITSLQWFGRNMVATSSDDGTVKLWD 209


RESULT   3
US-08-190-802A-30
; Sequence 30, Application US/08190802A
; Patent No. 5519003
;  GENERAL INFORMATION:
;    APPLICANT:  Mochly-Rosen, Daria
;    APPLICANT:  Ron, Dorit
;    TITLE OF INVENTION:  WD-40 - Derived Peptides and Uses
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-0850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/190,802A
;      FILING DATE:  01-FEB-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  8600-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  517 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  BETA TRCP, Fig. 13
```

```
US-08-190-802A-30

  Query Match             19.8%;  Score 623;  DB 1;  Length 517;
  Best Local Similarity   31.1%;  Pred. No. 2.6e-54;
  Matches  161;  Conservative   89;  Mismatches  214;  Indels   54;  Gaps   13;

Qy     60 VRSFSLGKKPCKVSEYTSTTGLVPCSATPT---TFGDLRAANGQGQQ---RRRITSVQPP 113
          :  ||   :|   ||        |     |   |    : |||       ::|  |
Db      1 MEGFSCSLQPPTASEREDCNRDEPPRKIITEKNTLRQTKLANGTSSMIVPKQRKLSANYE 60

Qy    114 TGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLP----KE 169
             :  :| |: ||  :::  :: ||      |  |:   ::|  |||||: ||
Db     61 KEKELCVKYFEQWSECDQVEFVEHLISRMCHYQHGHINTYLKPMLQRDFITALPARGLDH 120

Qy    170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKE--------EGIDEPLHIKRR 221
          :|  :||:|: | |  |   |: |  :  | :||::  :          |: |     ||
Db    121 IAENILSYLDAKSLCSAELVCKEWYRVTSDGMLWKKLIERMVRTDSLWRGLAE-----RR 175

Qy    222 K------VIKPGFIHSPWKSAY------IRQ--HRIDTNWRRGELKSPKV-LKGHDDHVI 266
                   ||    :|  | |      | |    |::||| |     ::  :      :
Db    176 GWGQYLFKNKPPDGKTPPNSFYRALYPKIIQDIETIESNWRCGRHSLQRIHCRSETSKGV 235

Qy    267 TCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLK 326
           |||:   :||||  |||:|:|   | :| | |:|||| |   |  : :||:|| | |::
Db    236 YCLQYDDQKIVSGLRDNTIKIWDKNTLECKRVLMGHTGSVLCLQYDERVIITGS-DSTVR 294

Qy    327 VWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL---HVLMGHV 383
          ||:  ||| ::||  |   |  :  :   :|: |:| :: |||: :   :    ||:||
Db    295 VWDVNTGEMLNTLIHHCEAVLHLRFNNGMMVTCSKDRSIAVWDMASATDITLRRVLVGHR 354

Qy    384 AAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDT 443
          |||  | :| : :|| : |  :|||:  |   : || ||   :  ||:    ||||| |
Db    355 AAVNVVDFDDKYIVSASGDRTIKVWNTSTCEFVRTLNGHKRGIACLQYRDRLVVSGSSDN 414

Qy    444 SIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTG---------Q 494
          :||:||:| | |:  | ||: |   :   :  :|||  |  :|:||:
Db    415 TIRLWDIECGACLRVLEGHEELVRCIRFDNKRIVSGAYDGKIKVWDLVAALDPRAPAGTL 474

Qy    495 CLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWD 532
          ||:||    :|   |  |||::  :::|| | |: :||
Db    475 CLRTLV---EHSGRVFRLQFDEFQIVSSSHDDTILIWD 509

RESULT   4
US-08-190-802A-32
; Sequence 32, Application US/08190802A
; Patent No. 5519003
;  GENERAL INFORMATION:
;    APPLICANT:  Mochly-Rosen, Daria
;    APPLICANT:  Ron, Dorit
;    TITLE OF INVENTION:  WD-40 - Derived Peptides and Uses
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-0850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/190,802A
;      FILING DATE:  01-FEB-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
```

```
;      REFERENCE/DOCKET NUMBER:  8600-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  779 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  CDC4 / CDC20 protein, Fig. 15
US-08-190-802A-32

  Query Match             19.7%;  Score 618.5;  DB 1;  Length 779;
  Best Local Similarity   29.6%;  Pred. No. 1.5e-53;
  Matches  152;  Conservative  105;  Mismatches  176;  Indels   81;  Gaps   16;

Qy    135 LDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYW- 193
          |  |: : : :::  :  :|:   :|| |: || |::| : ::|: :|:: :    : |
Db    247 LFRLVANMDRSELSDLGTLIKDNLKRDLITSLPFEISLKIFNYLQFEDIINSLGVSQNWN 306

Qy    194 RILAEDNLLWREKCKEEGIDEP-----LHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWR 248
          :|: :   ||::    |    |     |::|  :           : :::    |  ||
Db    307 KIIRKSTSLWKKLLISENFVSPKGFNSLNLKLSQKYPKLSQQDRLRLSFLENIFILKNW- 365

Qy    249 RGELKSPK------VLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGH 302
               :||       |:||   |||||||  | :::|:||  ::|: ::  | |  | ||
Db    366 ----YNPKFVPQRTTLRGHMTSVITCLQFEDNYVITGADDKMIRVYDSINKKFLLQLSGH 421

Qy    303 TGGVWS-SQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHE----KRVV 357
           ||||:       |::|||||||::||: : | | |   || |||||: : |    | :|
Db    422 DGGVWALKYAHGGILVSGSTDRTVRVWDIKKGCCTHVFEGHNSTVRCLDIVEYKNIKYIV 481

Qy    358 SGSRDATLRVWDI-----------------------ETGQCLHVLMGHVAAVRCVQYDGR 394
          :|||| || || :                       |    : || ||:|:|| |   |
Db    482 TGSRDNTLHVWKLPKESSVPDHGEEHDYPLVFHTPEENPYFVGVLRGHMASVRTVSGHGN 541

Qy    395 RVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIH----VVSGSLDTSIRVWDV 450
           ||||:||  : |||     ||: | |||:|:||  :|  |     :| |:||:||:||:
Db    542 IVVSGSYDNTLIVWDVAQMKCLYILSGHTDRIYSTIYD--HERKRCISASMDTTIRIWDL 599

Qy    451 E----TGNC----------------IHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDI 490
          |     | |                ::|| || :|   : | |  |||  || ::: ||
Db    600 ENIWNNGECSYATNSASPCAKILGAMYTLQGHTALVGLLRLSDKFLVSAAADGSIRGWD- 658

Qy    491 KTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGS 550
                :     : : ||:|    : | ::: |::    :::|::|: :   :  ::
Db    659 -ANDYSRKFSYHHTNLSAITTFYVSDNILVSGSEN-QFNIYNLRSGKLVHANILKDA--- 713

Qy    551 GGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDF 584
             :| :      || ||  ::|  :: | :|||
Db    714 -DQIWSVNFKGKTLVAAV-EKDG--QSFLEILDF 743

RESULT   5
US-08-283-917-3
; Sequence 3, Application US/08283917
; Patent No. 5849557
;  GENERAL INFORMATION:
;    APPLICANT:  ADACHI, HIDEKI
;    APPLICANT:  TSUJIMOTO, MASAFUMI
;    APPLICANT:  INOUE, KEIZO
;    APPLICANT:  ARAI, HIROYUKI
;    TITLE OF INVENTION:  OXIDIZED PHOPHOLIPID DEGRADING ENZYME
;    TITLE OF INVENTION:  AND GENE THEREOF
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER &
;      ADDRESSEE:  NEUSTADT,P.C.
```

```
;       STREET:  1755 S. Jefferson Davis Highway, Suite 400
;       CITY:  Arlington
;       STATE:  Virginia
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/283,917
;       FILING DATE:  03-AUG-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 209943/1993
;       FILING DATE:  03-AUG-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Oblon, No. 5849557man F.
;       REGISTRATION NUMBER:  24,618
;       REFERENCE/DOCKET NUMBER:  2292-030-0
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 413-3000
;       TELEFAX:  (703) 413-2220
;       TELEX:  248855 OPAT UR
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  409 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  Bos taurus
US-08-283-917-3

  Query Match             14.0%;  Score 439;  DB 2;  Length 409;
  Best Local Similarity   30.1%;  Pred. No. 7.2e-36;
  Matches  110;  Conservative   62;  Mismatches  133;  Indels   60;  Gaps   10;

Qy    227 GFIHSPWKSAYIRQHRI--------------DTNWRRGELKSPK----------VLKGHD 262
          | :   | |    | ::               :    |: : ||           | ||
Db     48 GLLEKKWTSVIRLQKKVMELESKLNEAKEEFTSGGPLGQKRDPKEWIPRPPEKYALSGHR 107

Qy    263 DHVITCL-QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVW--SSQMRDNIIISG 319
            |   :     : :|| |:| |:|||   ||   ||| |||  |   |      :: |
Db    108 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASC 167

Qy    320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377
          | | |:|:|: :  ||| |::||   |  : :  :   :|| ||| |:::|:::|| |:
Db    168 SADMTIKLWDFQGFECIRTMHGHDHNVSSVAIMPNGDHIVSASRDKTIKMWEVQTGYCVK 227

Qy    378 VLMGHVAAVRCVQ--YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---- 431
             ||   || |:   ||  : | : |  |:||   |: |   |: | : |  : :
Db    228 TFTGHREWVRMVRPNQDGTLIASCSNDQTVRVWVVATKECKAELREHEHVVECISWAPES 287

Qy    432 ------------------DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
                             |  ::||| | :|::||| || |: || || :   |:
Db    288 SYSSISEATGSETKKSGKPGPFLLSGSRDKTIKMWDVSTGMCLMTLVGHDNWVRGVLFHS 347

Qy    474 --NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVK 529
              ::|   | |:::|| |  :|::||   | |:  || | |:|   :|:| | | |||
Db    348 GGKFILSCADDKTLRVWDYKNKRCMKTL---NAHEHFVTSLDFHKTAPYVVTGSVDQTVK 404

Qy    530 LWDLK 534
          :|: :
Db    405 VWECR 409


RESULT   6
US-08-961-716-3
; Sequence 3, Application US/08961716
```

```
; Patent No. 5880272
;  GENERAL INFORMATION:
;     APPLICANT:  ADACHI, HIDEKI
;     APPLICANT:  TSUJIMOTO, MASAFUMI
;     APPLICANT:  INOUE, KEIZO
;     APPLICANT:  ARAI, HIROYUKI
;     TITLE OF INVENTION:  OXIDIZED PHOPHOLIPID DEGRADING ENZYME
;     TITLE OF INVENTION:  AND GENE THEREOF
;     NUMBER OF SEQUENCES:  31
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER &
;       ADDRESSEE:  NEUSTADT,P.C.
;       STREET:  1755 S. Jefferson Davis Highway, Suite 400
;       CITY:  Arlington
;       STATE:  Virginia
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/961,716
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/283,917
;       FILING DATE:  03-AUG-1994
;       APPLICATION NUMBER:  JP 209943/1993
;       FILING DATE:  03-AUG-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Oblon, No. 5880272man F.
;       REGISTRATION NUMBER:  24,618
;       REFERENCE/DOCKET NUMBER:  2292-030-0
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 413-3000
;       TELEFAX:  (703) 413-2220
;       TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  409 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  Bos taurus
US-08-961-716-3


  Query Match             14.0%;  Score 439;  DB 2;  Length 409;
  Best Local Similarity   30.1%;  Pred. No. 7.2e-36;
  Matches  110;  Conservative   62;  Mismatches  133;  Indels   60;  Gaps   10;

Qy    227 GFIHSPWKSAYIRQHRI--------------DTNWRRGELKSPK----------VLKGHD 262
          | :   | |    | ::               :    |: : ||           | ||
Db     48 GLLEKKWTSVIRLQKKVMELESKLNEAKEEFTSGGPLGQKRDPKEWIPRPPEKYALSGHR 107

Qy    263 DHVITCL-QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVW--SSQMRDNIIISG 319
            |   :     : :|| |:| |:|||   ||   ||| |||  |   |      :: |
Db    108 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASC 167

Qy    320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377
          | | |:|:|: :  ||| |::||   |  : :  :   :|| ||| |:::|:::|| |:
Db    168 SADMTIKLWDFQGFECIRTMHGHDHNVSSVAIMPNGDHIVSASRDKTIKMWEVQTGYCVK 227

Qy    378 VLMGHVAAVRCVQ--YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---- 431
             ||   || |:   ||  : | : |  |:||   |: |   |: | : |  : :
Db    228 TFTGHREWVRMVRPNQDGTLIASCSNDQTVRVWVVATKECKAELREHEHVVECISWAPES 287

Qy    432 ------------------DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
                             |  ::||| | :|::||| || |: || || :   |:
```

```
Db    288 SYSSISEATGSETKKSGKPGPFLLSGSRDKTIKMWDVSTGMCLMTLVGHDNWVRGVLFHS 347

Qy    474 --NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVK 529
              ::|   | |:::|| |  :|::||   | |:  || | |:|   :|:| | | |||
Db    348 GGKFILSCADDKTLRVWDYKNKRCMKTL---NAHEHFVTSLDFHKTAPYVVTGSVDQTVK 404

Qy    530 LWDLK 534
          :|: :
Db    405 VWECR 409


RESULT   7
US-08-283-917-9
; Sequence 9, Application US/08283917
; Patent No. 5849557
;  GENERAL INFORMATION:
;    APPLICANT:  ADACHI, HIDEKI
;    APPLICANT:  TSUJIMOTO, MASAFUMI
;    APPLICANT:  INOUE, KEIZO
;    APPLICANT:  ARAI, HIROYUKI
;    TITLE OF INVENTION:  OXIDIZED PHOPHOLIPID DEGRADING ENZYME
;    TITLE OF INVENTION:  AND GENE THEREOF
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER &
;      ADDRESSEE:  NEUSTADT,P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/283,917
;      FILING DATE:  03-AUG-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 209943/1993
;      FILING DATE:  03-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5849557man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  2292-030-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  410 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-283-917-9

  Query Match            14.0%;  Score 439;  DB 2;  Length 410;
  Best Local Similarity   30.1%;  Pred. No. 7.2e-36;
  Matches  110;  Conservative   62;  Mismatches  133;  Indels   60;  Gaps   10;

Qy    227 GFIHSPWKSAYIRQHRI--------------DTNWRRGELKSPK----------VLKGHD 262
          | :   | |    | ::               :    |: : ||           | ||
Db     49 GLLEKKWTSVIRLQKKVMELESKLNEAKEEFTSGGPLGQKRDPKEWIPRPPEKYALSGHR 108

Qy    263 DHVITCL-QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVW--SSQMRDNIIISG 319
            |   :     : :|| |:| |:|||   ||   ||| |||  |   |      :: |
Db    109 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASC 168

Qy    320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377
```

```
          | | |:|:|: :  ||| |::||   |  : :  :   :|| ||| |:::|:::|| |:
Db    169 SADMTIKLWDFQGFECIRTMHGHDHNVSSVAIMPNGDHIVSASRDKTIKMWEVQTGYCVK 228

Qy    378 VLMGHVAAVRCVQ--YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---- 431
             ||   || |:   ||  : | : |  |:||   |: |   |: | : |  : :
Db    229 TFTGHREWVRMVRPNQDGTLIASCSNDQTVRVWVVATKECKAELREHEHVVECISWAPES 288

Qy    432 ------------------DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
                             |  ::||| | :|::||| || |: || || :   |:
Db    289 SYSSISEATGSETKKSGKPGPFLLSGSRDKTIKMWDVSTGMCLMTLVGHDNWVRGVLFHS 348

Qy    474 --NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVK 529
              ::|   | |:::|| |  :|::||   | |:  || | |:|   :|:| | | |||
Db    349 GGKFILSCADDKTLRVWDYKNKRCMKTL---NAHEHFVTSLDFHKTAPYVVTGSVDQTVK 405

Qy    530 LWDLK 534
          :|: :
Db    406 VWECR 410


RESULT   8
US-08-961-716-9
; Sequence 9, Application US/08961716
; Patent No. 5880272
;  GENERAL INFORMATION:
;    APPLICANT:  ADACHI, HIDEKI
;    APPLICANT:  TSUJIMOTO, MASAFUMI
;    APPLICANT:  INOUE, KEIZO
;    APPLICANT:  ARAI, HIROYUKI
;    TITLE OF INVENTION:  OXIDIZED PHOPHOLIPID DEGRADING ENZYME
;    TITLE OF INVENTION:  AND GENE THEREOF
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER &
;      ADDRESSEE:  NEUSTADT,P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,716
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/283,917
;      FILING DATE:  03-AUG-1994
;      APPLICATION NUMBER:  JP 209943/1993
;      FILING DATE:  03-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5880272man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  2292-030-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  410 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-961-716-9

  Query Match            14.0%;  Score 439;  DB 2;  Length 410;
  Best Local Similarity   30.1%;  Pred. No. 7.2e-36;
```

```
  Matches  110;  Conservative   62;  Mismatches  133;  Indels   60;  Gaps   10;

Qy    227 GFIHSPWKSAYIRQHRI--------------DTNWRRGELKSPK----------VLKGHD 262
          | :   | |    | ::               :    |: : ||           | ||
Db     49 GLLEKKWTSVIRLQKKVMELESKLNEAKEEFTSGGPLGQKRDPKEWIPRPPEKYALSGHR 108

Qy    263 DHVITCL-QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVW--SSQMRDNIIISG 319
            |   :     : :|| |:| |:|||   ||   ||| |||  |   |      :: |
Db    109 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASC 168

Qy    320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377
          | | |:|:|: :  ||| |::||   |  : :  :   :|| ||| |:::|:::|| |:
Db    169 SADMTIKLWDFQGFECIRTMHGHDHNVSSVAIMPNGDHIVSASRDKTIKMWEVQTGYCVK 228

Qy    378 VLMGHVAAVRCVQ--YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---- 431
             ||   || |:   ||  : | : |  |:||   |: |   |: | : |  : :
Db    229 TFTGHREWVRMVRPNQDGTLIASCSNDQTVRVWVVATKECKAELREHEHVVECISWAPES 288

Qy    432 ------------------DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
                             |  ::||| | :|::||| || |: || || :   |:
Db    289 SYSSISEATGSETKKSGKPGPFLLSGSRDKTIKMWDVSTGMCLMTLVGHDNWVRGVLFHS 348

Qy    474 --NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVK 529
              ::|   | |:::|| |  :|::||   | |:  || | |:|   :|:| | | |||
Db    349 GGKFILSCADDKTLRVWDYKNKRCMKTL---NAHEHFVTSLDFHKTAPYVVTGSVDQTVK 405

Qy    530 LWDLK 534
          :|: :
Db    406 VWECR 410


RESULT   9
US-08-190-802A-51
; Sequence 51, Application US/08190802A
; Patent No. 5519003
;  GENERAL INFORMATION:
;    APPLICANT:  Mochly-Rosen, Daria
;    APPLICANT:  Ron, Dorit
;    TITLE OF INVENTION:  WD-40 - Derived Peptides and Uses
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-0850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/190,802A
;      FILING DATE:  01-FEB-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  8600-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  409 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
```

```
;       INDIVIDUAL ISOLATE:  LIS1 (human), Fig. 34
US-08-190-802A-51


  Query Match             13.5%;  Score 424.5;  DB 1;  Length 409;
  Best Local Similarity   29.9%;  Pred. No. 2.1e-34;
  Matches  109;  Conservative   62;  Mismatches  133;  Indels   61;  Gaps   11;

Qy    227 GFIHSPWKSAYIRQHRI--------------DTNWRRGELKSPK----------VLKGHD 262
          | :   | |    | ::               :    |: : ||           | ||
Db     49 GLLEKKWTSVIRLQKKVMELESKLNEAKEEFTSGGPLGQKRDPKEWIPRPPEKYALSGHR 108

Qy    263 DHVITCL-QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVW--SSQMRDNIIISG 319
            |   :     : :|| |:| |:|||   ||   ||| |||  |   |      :: |
Db    109 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASC 168

Qy    320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377
          | | |:|:|: :  ||| |::||   |  : :  :   :|| ||| |:::|:::|| |:
Db    169 SADMTIKLWDFQGFECIRTMHGHDHNVSSVAIMPNGDHIVSASRDKTIKMWEVQTGYCVK 228

Qy    378 VLMGHVAAVRCVQ--YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---- 431
             ||   || |:   ||  : | : |  |:||   |: |   |: | : |  : :
Db    229 TFTGHREWVRMVRPNQDGTLIASCSNDQTVRVWVVATKECKAELREHEHVVECISWAPES 288

Qy    432 ------------------DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
                             |  ::||| | : ::||| || |: || || :   |:
Db    289 SYSSISEATGSETKKSGKPGPFLLSGSRDKT-KMWDVSTGMCLMTLVGHDNWVRGVLFHS 347

Qy    474 --NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVK 529
              ::|   | |:::|| |  :|::||   | |:  || | |:|   :|:| | | |||
Db    348 GGKFILSCADDKTLRVWDYKNKRCMKTL---NAHEHFVTSLDFHKTAPYVVTGSVDQTVK 404

Qy    530 LWDLK 534
          :|: :
Db    405 VWECR 409


RESULT  10
US-08-899-578-7
; Sequence 7, Application US/08899578
; Patent No. 6087153
;  GENERAL INFORMATION:
;    APPLICANT:  Greenwald, Iva
;    APPLICANT:  Hubbard, E. Jane
;    TITLE OF INVENTION:  SEL-10 AND USES THEREOF
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/899,578
;      FILING DATE:  24-JUL-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/53200/JPW/AKC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 278-0525
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  212 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-899-578-7

  Query Match             13.0%;  Score 410;  DB 3;  Length 212;
  Best Local Similarity   34.9%;  Pred. No. 2.1e-33;
  Matches   96;  Conservative   41;  Mismatches   60;  Indels   78;  Gaps   10;

Qy    260 GHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWS-SQMRDNIIIS 318
          ||   |||||||  | :::|:||  ::|:            || ||||:       |::|
Db      1 GHMTSVITCLQFEDNYVITGADDKMIRVYD-----------GHDGGVWALKYAHGGILVS 49

Qy    319 GSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHE----KRVVSGSRDATLRVWDIETGQ 374
          ||||||::||:           || |||||: : |    | :|:  || || ||
Db     50 GSTDRTVRVWD-----------GHNSTVRCLDIVEYKNIKYIVTSGRDNTLHVW------ 92

Qy    375 CLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGI 434
                ||:|:|| |   |  ||||:||  : |||           |||:|:||  :|
Db     93 -----KGHMASVRTVSGHGNIVVSGSYDNTLIVWD-----------GHTDRIYSTIYD-- 134

Qy    435 H----VVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDI 490
          |     :| |:||:||:||           || :|   : | |  |||  || ::: ||
Db    135 HERKRCISASMDTTIRIWD-----------GHTALVGLLRLSDKFLVSAAADGSIRGWDH 183

Qy    491 KTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDD 525
           |            : ||:|    : | ::: |::
Db    184 HT------------NLSAITTFYVSDNILVSGSEN 206

RESULT  11
US-08-190-802A-52
; Sequence 52, Application US/08190802A
; Patent No. 5519003
;  GENERAL INFORMATION:
;    APPLICANT:  Mochly-Rosen, Daria
;    APPLICANT:  Ron, Dorit
;    TITLE OF INVENTION:  WD-40 - Derived Peptides and Uses
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-0850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/190,802A
;      FILING DATE:  01-FEB-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  8600-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  422 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
```

```
;      INDIVIDUAL ISOLATE:  MD6, Fig. 35
US-08-190-802A-52

  Query Match             12.0%;  Score 376.5;  DB 1;  Length 422;
  Best Local Similarity   25.6%;  Pred. No. 1.6e-29;
  Matches  117;  Conservative   75;  Mismatches  140;  Indels  125;  Gaps   18;

Qy    117 QEWLK----MFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELAL 172
          : ||      | |    :|   || ||      |::|:   :|   :|||: ||| ||:
Db      7 ETWLDNISVTFLSLMDLQKNETLDHLISLSGAVQLRHLSNNLETLLKRDFLKLLPLELSF 66

Qy    173 YVLSFLEPKDLLQAAQTC-----RYWRILAEDNLLWREKCK------EEGIDEPLHIKRR 221
          |:| :|:|: ||    ||     :  ::::    :|:  ||      :: : : ||
Db     67 YLLKWLDPQTLL----TCCLVSKQRNKVISACTEVWQTACKNLGWQIDDSVQDSLH---- 118

Qy    222 KVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSD 281
                     ||  |              ||:   :|  :||
Db    119 -----------WKKVY--------------LKAILRMKQLEDH----------------- 136

Qy    282 DNTLKVWSAVTGKCLRT--LVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTL 339
                      :   |  |:||:  |::   :| :: :|| | : |:|:  ||:|:: :
Db    137 ------------EAFETSSLIGHSARVYALYYKDGLLCTGSDDLSAKLWDVSTGQCVYGI 184

Qy    340 YGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRR--VV 397
            ||     :   |:::|:|| | |:  |:  :|       ||  ||  | |      :|
Db    185 QTHTCA--AVKFDEQKLVTGSFDNTVACWEWSSGARTQHFRGHTGAVFSVDYSDELDILV 242

Qy    398 SGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIH 457
          ||: || ||||     |||:|| |||  |                | : :   :  : :|
Db    243 SGSADFAVKVWALSAGTCLNTLTGHTEWV----------------TKVVLQKCKVKSLLH 286

Qy    458 TLTGHQSLTSGMELKDNILVSGNADS-TVKIWDIK---TGQCLQTLQGPNKHQSAVTCLQ 513
          :              |  ||:|  ||   :||| |      :||:||   :  :    |||
Db    287 S------------PGDYILLS--ADKYEIKIWPIGREINCKCLKTL---SVSEDRSICLQ 329

Qy    514 ----FNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLE 546
              |:  ::: ||  |  : ||  : : :| : | |
Db    330 PRLHFDGKYIVCSSALGLYQ-WDFASYDILRVIKTPE 365

RESULT  12
US-08-190-802A-66
; Sequence 66, Application US/08190802A
; Patent No. 5519003
;  GENERAL INFORMATION:
;    APPLICANT:  Mochly-Rosen, Daria
;    APPLICANT:  Ron, Dorit
;    TITLE OF INVENTION:  WD-40 - Derived Peptides and Uses
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-0850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/190,802A
;      FILING DATE:  01-FEB-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  8600-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
```

```
;      TELEFAX:  (415) 324-0960
;   INFORMATION FOR SEQ ID NO:  66:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  514 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       INDIVIDUAL ISOLATE:  YCW2 PROTEIN, Fig. 49
US-08-190-802A-66


  Query Match             11.6%;  Score 364.5;  DB 1;  Length 514;
  Best Local Similarity   23.0%;  Pred. No. 3.6e-28;
  Matches  119;  Conservative  116;  Mismatches  153;  Indels  129;  Gaps   26;

Qy    155 EPQFQRDFISLLPKELALYVLSF--LEPKDLL-------------QAAQTCRYWRILAED 199
          | |  |: ::::||:|    : |  |:  | :             |  :        ::|
Db     15 EAQLPRE-VAIIPKDLPNVSIKFQALDTGDNVGGALRVPGAISEKQLEELLNQLNGTSDD 73

Qy    200 NLLWREKCKEEG-----------IDEPLHIKRRKVIKPGF---------IHSPWKSAYIR 239
           : :   |  :|           | : |:     :||||:         :::|     ::
Db     74 PVPYTFSCTIQGKKASDPVKTIDITDNLY---SSLIKPGYNSTEDQITLLYTPRAVFKVK 130

Qy    240 QHRIDTNWRRGELKSPKVLKGHDDHVITCLQFC---GNRIVSGSDDNTLKVWSAVTGKCL 296
                       :|   : ||    | |  |     :|:|:|: ||| ::|   |   :
Db    131 P----------VTRSSSAIAGHGS-TILCSAFAPHTSSRMVTGAGDNTARIWDCDTQTPM 179

Qy    297 RTLVGHTGGV----WSSQMRDNIIISGSTDRTLKVWNAETGECI-HTLYGHTSTVRCM-- 349
           || ||   |    ||      :| :|| | |:::|: ::|:|:   | ||:  :  :
Db    180 HTLKGHYNWVLCVSWSPD--GEVIATGSMDNTIRLWDPKSGQCLGDALRGHSKWITSLSW 237

Qy    350 ---HL----HEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVV-SGAY 401
             ||     : |: | |:| |:::||  :  | : : ||  :| ||:: |: :: ||::
Db    238 EPIHLVKPGSKPRLASSSKDGTIKIWDTVSRVCQYTMSGHTNSVSCVKWGGQGLLYSGSH 297

Qy    402 DFMVKVWDPETE-TCLHTLQGHTNRV--------YSLQ---FDGIH-------------- 435
          |  |:|||  ::  |:: |: | : |        |:|:   ||  |
Db    298 DRTVRVWDINSQGRCINILKSHAHWVNHLSLSTDYALRIGAFD--HTGKKPSTPEEAQKK 355

Qy    436 ------------------VVSGSLDTSIRVWD-VETGNCIHTLTGHQSLTSGMELKDN-- 474
                            :|: | | :: :|: :::   |  :|||| | : :    :
Db    356 ALENYEKICKKNGNSEEMMVTASDDYTMFLWNPLKSTKPIARMTGHQKLVNHVAFSPDGR 415

Qy    475 ILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVKLWD 532
           :|| : |:::|:|| : |: : | :|   | ::|  : :: :   ::: | | |:|:||
Db    416 YIVSASFDNSIKLWDGRDGKFISTFRG---HIASVYQVAWSSDCRLLVSCSKDTTLKVWD 472

Qy    533 LKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVG 569
          ::| :   :|  :::     | | :   : | ||: |
Db    473 VRTRKLSVDLPGIKT--KLYVDWSV---DGKRVCSGG 504


RESULT  13
US-08-188-582-5
; Sequence 5, Application US/08188582
; Patent No. 5534410
;  GENERAL INFORMATION:
;    APPLICANT:  Tjian, Robert
;    APPLICANT:  Comai, Lucio
;    APPLICANT:  Dynlact, Brian D.
;    APPLICANT:  Hoey, Timothy
;    APPLICANT:  Ruppert, Siegfried
;    APPLICANT:  Tanese, Naoko
;    APPLICANT:  Wang, Edith
;    APPLICANT:  Weinzierl, Robert O.J.
;    TITLE OF INVENTION:  TATA-BINDING PROTEIN ASSOCIATED FACTORS,
;    TITLE OF INVENTION:  NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
```

```
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/188,582
;      FILING DATE:  28-JAN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  A-57650-2/AJT/RAO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  704 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-188-582-5


  Query Match             10.6%;  Score 332;  DB 1;  Length 704;
  Best Local Similarity   29.0%;  Pred. No. 1.2e-24;
  Matches   88;  Conservative   61;  Mismatches  126;  Indels   28;  Gaps   13;

Qy    257 VLKGHDDHVITCLQFC--GNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDN 314
          ||  |    :|| :       :  |  |:::::||    | ||||        :  :|:
Db    373 VLNSHQG--VTCAEISDDSTMLACGFGDSSVRIWSLTPAK-LRTLKD------ADSLRE- 422

Qy    315 IIISGSTDRTLKVWNAETGECIHTLYGHTSTV-RCMHLHEKR-VVSGSRDATLRVWDIET 372
           :   | |  ::: :  :||   :| |||  | ||    |   ::| | |:|:|:| : |
Db    423 -LDKESADINVRMLDDRSGEVTRSLMGHTGPVYRCAFAPEMNLLLSCSEDSTIRLWSLLT 481

Qy    373 GQCLHVLMGHVAAVRCVQY--DGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQ 430
            |:    |||  |  |::   |   || :||   ::|  ::   |    || : |  :|
Db    482 WSCVVTYRGHVYPVWDVRFAPHGYYFVSCSYDKTARLWATDSNQALRVFVGHLSDVDCVQ 541

Qy    431 F--DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELK--DNILVSGNADSTVK 486
          |  :  :| :|| | ::|:||  ||  :  :|||:   | :        | ||: |  :
Db    542 FHPNSNYVATGSSDRTVRLWDNMTGQSVRLMTGHKGSVSSLAFSACGRYLASGSVDHNII 601

Qy    487 IWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSD--DGTVKLWDLK--TGEFIRNL 542
          |||:  |  : ||    :| | || : |:::  : ::   |  : |||    | ::| |
Db    602 IWDLSNGSLVTTLL---RHTSTVTTITFSRDGTVLAAAGLDNNLTLWDFHKVTEDYISNH 658

Qy    543 VTL 545
          :|:
Db    659 ITV 661


RESULT  14
US-08-646-715-5
; Sequence 5, Application US/08646715
; Patent No. 5637686
;  GENERAL INFORMATION:
;    APPLICANT:  Tjian, Robert
;    APPLICANT:  Comai, Lucio
;    APPLICANT:  Dynlact, Brian D.
;    APPLICANT:  Hoey, Timothy
;    APPLICANT:  Ruppert, Siegfried
;    APPLICANT:  Tanese, Naoko
;    APPLICANT:  Wang, Edith
;    APPLICANT:  Weinzierl, Robert O.J.
```

```
;    TITLE OF INVENTION:  TATA-BINDING PROTEIN ASSOCIATED FACTORS,
;    TITLE OF INVENTION:  NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/646,715
;      FILING DATE:  09-MAY-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/188,582
;      FILING DATE:  28-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  A-57650-2/AJT/RAO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  704 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-646-715-5


  Query Match             10.6%;  Score 332;  DB 1;  Length 704;
  Best Local Similarity   29.0%;  Pred. No. 1.2e-24;
  Matches   88;  Conservative   61;  Mismatches  126;  Indels   28;  Gaps   13;

Qy     257 VLKGHDDHVITCLQFC--GNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDN 314
           ||  |    :|| :       :  |  |:::::||    | ||||        :  :|:
Db     373 VLNSHQG--VTCAEISDDSTMLACGFGDSSVRIWSLTPAK-LRTLKD------ADSLRE- 422

Qy     315 IIISGSTDRTLKVWNAETGECIHTLYGHTSTV-RCMHLHEKR-VVSGSRDATLRVWDIET 372
            :   | |  ::: :  :||   :| |||  | ||    |   ::| | |:|:|:| : |
Db     423 -LDKESADINVRMLDDRSGEVTRSLMGHTGPVYRCAFAPEMNLLLSCSEDSTIRLWSLLT 481

Qy     373 GQCLHVLMGHVAAVRCVQY--DGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQ 430
             |:    |||  |  |::   |   || :||   ::|  ::   |    || : |  :|
Db     482 WSCVVTYRGHVYPVWDVRFAPHGYYFVSCSYDKTARLWATDSNQALRVFVGHLSDVDCVQ 541

Qy     431 F--DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELK--DNILVSGNADSTVK 486
           |  :  :| :|| | ::|:||  ||  :  :|||:   | :        | ||: |  :
Db     542 FHPNSNYVATGSSDRTVRLWDNMTGQSVRLMTGHKGSVSSLAFSACGRYLASGSVDHNII 601

Qy     487 IWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSD--DGTVKLWDLK--TGEFIRNL 542
           |||:  |  : ||    :| | || : |:::  : ::   |  : |||    | ::| |
Db     602 IWDLSNGSLVTTLL---RHTSTVTTITFSRDGTVLAAAGLDNNLTLWDFHKVTEDYISNH 658

Qy     543 VTL 545
           :|:
Db     659 ITV 661


RESULT  15
US-08-190-802A-62
; Sequence 62, Application US/08190802A
; Patent No. 5519003
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Mochly-Rosen, Daria
;    APPLICANT:  Ron, Dorit
;    TITLE OF INVENTION:  WD-40 - Derived Peptides and Uses
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-0850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/190,802A
;      FILING DATE:  01-FEB-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  8600-0139
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  62:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  704 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  TRANSCRIPTION FACTOR TIIF, Fig. 45
US-08-190-802A-62


  Query Match             10.4%;  Score 326;  DB 1;  Length 704;
  Best Local Similarity   28.7%;  Pred. No. 4.7e-24;
  Matches   87;  Conservative   63;  Mismatches  125;  Indels   28;  Gaps   13;

Qy     257 VLKGHDDHVITCLQFC--GNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDN 314
           ||  |    :|| :       :  |  |:::::|| :|   :|||        :  :|:
Db     373 VLNSHQG--VTCAEISDDSTMLACGFGDSSVRIWS-LTPANVRTLKD------ADSLRE- 422

Qy     315 IIISGSTDRTLKVWNAETGECIHTLYGHTSTV-RCMHLHEKR-VVSGSRDATLRVWDIET 372
            :   | |  ::: :  :||   :| |||  | ||    |   ::| | |:|:|:| : |
Db     423 -LDKESADINVRMLDDRSGEVTRSLMGHTGPVYRCAFAPEMNLLLSCSEDSTIRLWSLLT 481

Qy     373 GQCLHVLMGHVAAVRCVQY--DGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQ 430
             |:    |||  |  |::   |   || :||   ::|  ::   |    || : |  :|
Db     482 WSCVVTYRGHVYPVWDVRFAPHGYYFVSCSYDKTARLWATDSNQALRVFVGHLSDVDCVQ 541

Qy     431 F--DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELK--DNILVSGNADSTVK 486
           |  :  :| :|| | ::|:||  ||  :  :|||:   | :        | ||: |  :
Db     542 FHPNSNYVATGSSDRTVRLWDNMTGQSVRLMTGHKGSVSSLAFSACGRYLASGSVDHNII 601

Qy     487 IWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSD--DGTVKLWDLK--TGEFIRNL 542
           |||:  |  : ||    :| | || : |:::  : ::   |  : |||    | ::| |
Db     602 IWDLSNGSLVTTLL---RHTSTVTTITFSRDGTVLAAAGLDNNLTLWDFHKVTEDYISNH 658

Qy     543 VTL 545
           :|:
Db     659 ITV 661


Search completed: September 25, 2001, 14:41:34
Job time: 154 sec
```

```
                    GenCore version 4.5
           Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         September 25, 2001, 14:39:35 ; Search time 17.58 Seconds
                                              (without alignments)
                                              2552.151 Million cell updates/sec

Title:          US-09-328-877A-8
Perfect score:  3143
Sequence:       1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                 1:  pir1:*
                 2:  pir2:*
                 3:  pir3:*
                 4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result              Query
  No.     Score     Match Length DB  ID                         Description
-------------------------------------------------------------------------------
    1      1238     39.4    579  2  T22703                    hypothetical prote
    2       680     21.6    605  2  T38932                    probable sulfur me
    3       650     20.7    640  2  S49932                    MET30 protein - ye
    4       641     20.4    703  2  T43557                    F-box/WD-repeat pr
    5     639.5     20.3    701  2  T16607                    hypothetical prote
    6     634.5     20.2    518  2  B48088                    beta-transducin re
    7       620     19.7    775  2  T45136                    WD repeat protein
    8     618.5     19.7    779  2  S56245                    cell division cont
    9     582.5     18.5   1356  2  T18521                    beta transducin-li
   10     576.5     18.3    650  2  T46660                    sulfur controller-
   11       533     17.0    506  2  T50211                    WD-repeat protein
   12     446.5     14.2   1146  2  A55532                    myosin-heavy-chain
   13       439     14.0    410  2  S48052                    platelet-activatin
   14     436.5     13.9    409  2  S36113                    LIS-1 protein - hu
   15     410.5     13.1    317  2  T46032                    WD-40 repeat regul
   16       403     12.8    267  2  S62507                    hypothetical trp-a
   17       396     12.6    376  2  T19266                    hypothetical prote
   18     392.5     12.5    333  2  G85034                    probable WD-repeat
   19     378.5     12.0    502  2  T41148                    trp-asp repeat con
   20     376.5     12.0    422  2  A56640                    CDC4 repeat unit-c
   21       376     12.0    501  2  T27513                    hypothetical prote
   22       365     11.6    515  2  S19487                    hypothetical prote
   23       361     11.5    395  2  T23317                    hypothetical prote
   24       360     11.5   1693  2  S76086                    beta transducin-li
   25     353.5     11.2    714  2  S56893                    hypothetical prote
   26       336     10.7    777  2  T41075                    hypothetical WD-re
   27     333.5     10.6    494  2  T19550                    hypothetical prote
   28       330     10.5    486  2  S49820                    PRL1 protein - Ara
   29       326     10.4    704  2  S33263                    transcription init
   30     324.5     10.3    651  2  T50289                    WD repeat protein
   31       322     10.2    504  2  T50983                    probable pleiotrop
   32     318.5     10.1    713  2  JN0133                    WD-40 repeat regul
   33       318     10.1    451  2  S65162                    hypothetical prote
   34       317     10.1    369  2  T40905                    WD repeat protein
   35     315.5     10.0    612  2  T08602                    protein TipD - sli
   36     314.5     10.0    431  2  S49821                    PRL2 protein - Ara
   37       314     10.0   1019  3  JC7538                    neuronal different
   38       312      9.9    659  2  S38108                    hypothetical prote
   39     311.5      9.9   1888  2  T14273                    zinc finger protei
   40       311      9.9    314  2  T43299                    probable GTP-bindi
   41       311      9.9    473  2  T33805                    hypothetical prote
   42     310.5      9.9    876  2  T51507                    WD40-repeat protei
   43       310      9.9    614  2  S58306                    WD-40 repeat regul
   44       306      9.7    798  2  S34023                    TATA box-binding p
   45     304.5      9.7    380  2  T40283                    beta-transducin -
```

ALIGNMENTS

```
RESULT   1
T22703
hypothetical protein F55B12.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22703
R;Sims, M.
submitted to the EMBL Data Library, September 1996
A;Reference number: Z19602
A;Accession: T22703
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-579 <WIL>
A;Cross-references: EMBL:Z79757; PIDN:CAB02129.1; GSPDB:GN00023; CESP:F55B12.3
A;Experimental source: clone F55B12
C;Genetics:
A;Gene: CESP:F55B12.3
A;Map position: 5
A;Introns: 22/3; 45/2; 77/3; 122/3; 171/2; 211/2; 342/2; 417/3; 478/3; 526/1

  Query Match             39.4%;  Score 1238;  DB 2;  Length 579;
  Best Local Similarity   45.6%;  Pred. No. 9.8e-90;
  Matches  262;  Conservative   85;  Mismatches  162;  Indels   66;  Gaps   11;

Qy     55 DHGSEVRSFSLGKKPC----KVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSV 110
          |:| |  |:| |        |:|        :  ||:|:   ||
Db     23 DNGEE-SSYSNGSSSSYNADKLSSSRPLQHKLDLSASPSRNNDLNPR------------- 68

Qy    111 QPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKEL 170
               ::  : :|:  |  |:: |   |:     | :: :  :||| |||||:| || ||
Db     69 -----VEHLIALFKDLSSAEQMDAFTRLLQESNMTNIRQLRAIIEPHFQRDFLSCLPVEL 123

Qy    171 ALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEE-------------------- 210
           : :|  |   |||: ||  : |::::| : :|:    ||
Db    124 GMKILHNLTGYDLLKVAQVSKNWKLISEIDKIWKSLGVEEFKHHPDPTDRVTGAWQGTAI 183

Qy    211 --GIDEPLHI--------KRRKVIKPG--FIHSPWKSAYIRQHRIDTNWRRGELKSPKVL 258
            |:  | ||        :  |: | |  |  :  || |:|  :|: ||    :    ||
Db    184 AAGVTIPDHIQPCDLNVHRFLKLQKFGDIFERAADKSRYLRADKIEKNWNANPIMGSAVL 243

Qy    259 KGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRD--NII 316
          :||:||||||:|   : :|:||||||||||    |: : ||||||||||:||:      |
Db    244 RGHEDHVITCMQIHDDVLVTGSDDNTLKVWCIDKGEVMYTLVGHTGGVWTSQISQCGRYI 303

Qy    317 ISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL 376
          :|||||||:|||:   |  :||| ||||||||| :    :|:|||| ||||||:|:|: |
Db    304 VSGSTDRTVKVWSTVDGSLLHTLQGHTSTVRCMAMAGSILVTGSRDTTLRVWDVESGRHL 363

Qy    377 HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIH- 435
            | || |||||||:||  |||| ||| ||:|:  |  |: || || |||||| |:
Db    364 ATLHGHHAAVRCVQFDGTTVVSGGYDFTVKIWNAHTGRCIRTLTGHNNRVYSLLFESERS 423
```

```
Qy    436 -VVSGSLDTSIRVWDV---ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIK 491
           | ||||||||||||    |   |:  | || ||||||:|: ||||| |||| |::|||
Db    424 IVCSGSLDTSIRVWDFTRPEGQECVALLQGHTSLTSGMQLRGNILVSCNADSHVRVWDIH 483

Qy    492 TGQCLQTLQGPNKHQSAVTCLQ-FNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGS 550
           | |:  | |   |:||:| || | :| | ||||||||||||:: |  ||:||||:|||:
Db    484 EGTCVHMLSG---HRSAITSLQWFGRNMVATSSDDGTVKLWDIERGALIRDLVTLDSGGN 540

Qy    551 GGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFD 585
          || :||: :::| | ||||||| |||||:::||||
Db    541 GGCIWRLCSTSTMLACAVGSRNNTEETKVILLDFD 575

RESULT   2
T38932
probable sulfur metabolite control protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
C;Accession: T38932
R;Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1997
A;Reference number: Z21818
A;Accession: T38932
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-605 <BAD>
A;Cross-references: EMBL:Z94864; PIDN:CAB08168.1; GSPDB:GN00066; SPDB:SPAC57A10.05c
A;Experimental source: strain 972h-; cosmid c57A10
C;Genetics:
A;Gene: SPDB:SPAC57A10.05c
A;Map position: 1
C;Superfamily: unassigned WD repeat proteins; WD repeat homology


  Query Match             21.6%;  Score 680;  DB 2;  Length 605;
  Best Local Similarity   29.8%;  Pred. No. 1.1e-45;
  Matches  164;  Conservative   93;  Mismatches  181;  Indels  112;  Gaps   17;

Qy     51 KRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSV 110
          |:: |    | :|::       | :   :||                    : |:|: :|
Db     31 KKEGDQSISVSAFNIS------SMHNELSGL------------------SEKSRQRVEAV 66

Qy    111 QPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKEL 170
                  |   |   |  |: |||  ::::|  : :      ::   : ||:|||| |:
Db     67 --------W-AAFSEASCSERKLALQGILNNCSSSLLSFASSTLDSLVRLDFLSLLPVEI 117

Qy    171 ALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKE--------------------- 209
          :  :||||: : | ||||  ::|: ||:|:::|   |::
Db    118 SFRILSFLDARSLCQAAQVSKHWKELADDDVIWHRMCEQHINRKCEKCGWGLPLLERNTL 177

Qy    210 ----------------EGID---EPLHIKRRKV--------------IKP---------- 226
                           |:|   |   :|: |:              :||
Db    178 YAAKASIQKRYERLTKRGVDQAHESSPVKKAKLDDYPTSSNEETISSVKPPSPNSDSKFF 237

Qy    227 -GFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTL 285
            |   |||  |  : |:: ||| |  :   || || | |: |||   | : ||| | |:
Db    238 LPFKTRPWKEVYAERCRVECNWRHGRCRQ-VVLSGHSDGVM-CLQLVRNILASGSYDATI 295

Qy    286 KVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTST 345
          ::|:  | : :  | ||: ||   |     :|||| |:|:::||  | |||  |:||| :
Db    296 RLWNLATFQQVALLEGHSSGVTCLQFDQCKLISGSMDKTIRIWNYRTSECISILHGHTDS 355

Qy    346 VRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQY--DGRRVVSGAYDF 403
          | |:      :|||| | |:::|    |: :  | ||   |  |:   |   |:||: |
Db    356 VLCLTFDSTLLVSGSADCTVKLWHFSGGKRI-TLRGHTGPVNSVRIIRDRGLVLSGSDDS 414

Qy    404 MVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQ 463
           :|:|  || |||||   |   | ||      : | ||| :|: ||:|   |:||| ||
Db    415 TIKIWSLETNTCLHTFSAHIGPVQSLALADSRLFSCSLDGTIKQWDIEKKKCVHTLFGH- 473

Qy    464 SLTSGMEL-KDNI-LVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVIT 521
           :    |:  |:: |:||  |  ||:|:    :|: ||:    |   || :      |::
Db    474 -IEGVWEIAADHLRLISGAHDGVVKVWE--ACECVHTLK---NHSEPVTSVALGDCEVVS 527

Qy    522 SSDDGTVKLW 531
           |:|| : ||
Db    528 GSEDGKIYLW 537

RESULT   3
S49932
MET30 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YI9905.02; protein YIL046w
C;Species: Saccharomyces cerevisiae
C;Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 26-May-2000
C;Accession: S49932; S43750
R;Odell, C.; Bowman, S.
submitted to the EMBL Data Library, December 1994
A;Reference number: S49931
A;Accession: S49932
A;Molecule type: DNA
A;Residues: 1-640 <ODE>
A;Cross-references: GB:Z47047; EMBL:Z46861; NID:g603997; PID:g763300; MIPS:YIL046w
R;Thomas, D.; Cherest, H.; Barbey, R.; Surdin-Kerjan, Y.
submitted to the EMBL Data Library, December 1993
A;Reference number: S43750
A;Accession: S43750
A;Molecule type: DNA
A;Residues: 1-60,'I',62-640 <THO>
A;Cross-references: EMBL:L26505; NID:g432493; PID:g432494
C;Genetics:
A;Gene: SGD:MET30
A;Cross-references: SGD:S0001308; MIPS:YIL046w
A;Map position: 9L
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
F;298-329/Domain: WD repeat homology <WD1>
F;338-369/Domain: WD repeat homology <WD2>
F;374-409/Domain: WD repeat homology <WD3>
F;417-450/Domain: WD repeat homology <WD4>


  Query Match             20.7%;  Score 650;  DB 2;  Length 640;
  Best Local Similarity   26.4%;  Pred. No. 2.9e-43;
  Matches  163;  Conservative  114;  Mismatches  197;  Indels  144;  Gaps   18;

Qy      7 PTLNHGLVPVDLKSAKEPLPHQTVMKIFSISIIAQGLPFCRRRMKRKLDHGSEVRSFSLG 66
          |: :  |   |  :  :|||     |            || |       |  ::: ||
Db     68 PSSSPDLATNDSGTRVQPLPEYNFTK------------FCYR-------HNPDIQ-FSPT 107

Qy     67 KKPC------KVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWL 120
             |      :  |  :    :|                          :|  : :   :
Db    108 HTACYKQDLKRTQEINANIAKLP--------------------------LQEQSDIHHII 141

Qy    121 KMFQSWSGPEKLLALDELID-SCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLE 179
            : : :   : | || ::  || | |: ::  ::    : ||||:||:||:| :||:|:
Db    142 SKYSNSNDKIRKLILDGILSTSCFP-QLSYISSLVTHMIKIDFISILPQELSLKILSYLD 200

Qy    180 PKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEP----------LHIKRRKVIKPGFI 229
           : |  | : || |: ||:|: :|   | |: ||            ||:|| :: :
Db    201 CQSLCNATRVCRKWQKLADDDRVWYHMC-EQHIDRKCPNCGWGLPLLHMKRARIQQNSTG 259

Qy    230 HS-----------PWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVS 278
           |           |||  |  : ::::|||:|  :  :  ||| | |:| |||    : :
Db    260 SSSNADIQTQTTRPWKVIYRERFKVESNWRKGHCRIQE-FKGHMDGVLT-LQFNYRLLFT 317

Qy    279 GSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHT 338
          || |:|: :|   ||| :| | ||: || :    |  :|:|| |:|::|||  ||||| |
Db    318 GSYDSTIGIWDLFTGKLIRRLSGHSDGVKTLYFDDRKLITGSLDKTIRVWNYITGECIST 377

Qy    339 LYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGR--RV 396
            ||: :|  :  ::| :|||| | |::|| :|:  | : | ||   | ||:   :
Db    378 YRGHSDSVLSVDSYQKVIVSGSADKTVKVWHVESRTC-YTLRGHTEWVNCVKLHPKSFSC 436

Qy    397 VSGAYDFMVKVWDPETETCLHTLQGHTNRVYSL--------------------------- 429
           | : |  :::||  | :||   :||  :|  :
```

```
Db     437 FSCSDDTTIRMWDIRTNSCLKVFRGHVGQVQKIIPLTIKDVENLATDNTSDGSSPQDDPT 496

Qy     430 QFDG----------------------IHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTS 467
             ||                       |::|  || :|::|||:|| || |  ||
Db     497 MTDGADESDTPSNEQETVLDENIPYPTHLLSCGLDNTIKLWDVKTGKCIRTQFGHVEGVW 556

Qy     468 GMELKDNILVSGNADSTVKIWDIKTGQCLQT-----LQGPNKHQ---------SAVTCLQ 513
            :   :  ::||: | ::|:||:::|:|: |     ||   :|          : : |:
Db     557 DIAADNFRIISGSHDGSIKVWDLQSGKCMHTFNGRRLQRETQHTQTQSLGDKVAPIACVC 616

Qy     514 FNKNFVITSSDDGTVKLW 531
              :   :  : | ||::
Db     617 IGDSECFSGDEFGCVKMY 634

RESULT   4
T43557
F-box/WD-repeat protein pop2 - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: proteolysis factor sud1p
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T43557; T38794; T43798
R;Wolf, D.A.; Jackson, P.K.
submitted to the EMBL Data Library, December 1997
A;Description: Fission yeast pop2 encodes a novel F-box/WD-repeat protein involved in th
A;Reference number: Z22576
A;Accession: T43557
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-703 <WOL>
A;Cross-references: EMBL:AF038867; PIDN:AAB95480.1
A;Experimental source: strain h- 972
R;Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z21812
A;Accession: T38794
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-703 <GEN>
A;Cross-references: EMBL:Z98602; PIDN:CAB11275.1; GSPDB:GN00066; SPDB:SPAC4D7.03
A;Experimental source: strain 972h-; cosmid c4D7
R;Jallepalli, P.V.; Tien, D.; Kelly, T.J.
Proc. Natl. Acad. Sci. U.S.A. 95, 8159-8164, 1998
A;Title: Sud1+ targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1 proteins f
A;Reference number: Z22686; MUID:98318628
A;Accession: T43798
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-703 <JAL>
A;Cross-references: EMBL:AF064515; NID:g3293382; PIDN:AAC39496.1; PID:g3293383
C;Genetics:
A;Gene: pop2; sud1; SPAC4D7.03
A;Map position: 1
C;Function:
A;Description: required to prevent spontaneous re-replication


  Query Match             20.4%;  Score 641;  DB 2;  Length 703;
  Best Local Similarity   27.3%;  Pred. No. 1.7e-42;
  Matches  178;  Conservative  111;  Mismatches  248;  Indels  116;  Gaps   17;

Qy       9 LNHGLVPVDLKSAKEPL-----------PHQTVMKIFSISIIAQ--------GLP----F 45
           | |   |  |:   | :           |  :|   ||||:: |         ||
Db      85 LKHSCTPSFLRKFNESIENVSYKCLDHSPPDSVPGDFSISLVPQRNFLYSHSSLPPKIIS 144

Qy      46 CRRRMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRR 105
             |  : |||:     | :    |   :  | : |  | |       ||   | |   :
Db     145 IDRNNRIKLDNSISSNSDNFPPSPKVDTSNTVSPGSKPISE------DLEDLNLQSIVQ- 197

Qy     106 RITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISL 165
             |    | |:| :              |  :|: ||    ::  ::   ||  ::| :|
Db     198 --TFEDLPEGIQSY--------------AFFQLLRSCNRQSMRLLLNECEPLLKKDILSN 241
```

```
Qy     166 LPKELALYVLSFLEPKDLLQAAQTCRYW-RILAEDNLLWREKCKEEGI---DEPLHIKRR 221
           ||  :   :|  |:    |        | |||      |:      |    :
Db     242 LPFSIVQSILLNLDIHSFLSCRLVSPTWNRILDVHTSYWKHMFSLFGFQINENDWKYANP 301

Qy     222 KVIKPGFIHSPWKS-----AYIRQHRIDTNWRRGELKSPKVLKGH-------DDHVITCL 269
            : :| |:|:   |        ::| ::    |     | :   |        : :|| |
Db     302 NLNRPPFLHNDQISDDYFPEIFKRHFLN----RKRWLFPSIPPSHLSFPIHVPNFMITSL 357

Qy     270 QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWN 329
               :||:: |   |::: :|:||     | ||  |||: :: :| ::||| |:|::|||
Db     358 LLHKDRIITTSGSGTIQIHNAITGVLEARLEGHKEGVWAVKIHENTLVSGSIDKTVRVWN 417

Qy     330 AETGECIHTLYGHTSTVRCMHL----------------HEKRVVSGSRDATLRVWDI--- 370
            |  :| |   || | :||: :                 :  :|||||| ||||| :
Db     418 IEKAKCTHIFRGHISIIRCLEILVPSRLIRHGVEIVEPDQPYIVSGSRDHTLRVWKLPKN 477

Qy     371 -----------------ETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETE 413
                            :    :| |:||  :|| :   |  :|||:||  :::|   |
Db     478 TDPPYLPDNTNSIDRWEKNPYFVHTLIGHTDSVRTISGYGDILVSGSYDSSIRIWRVSTG 537

Qy     414 TCLHTLQGHTNRVYSLQFDGIH--VVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMEL 471
            ||: |:||: |:||: ::      :|||:| ||||||: || | : | || :  : : :
Db     538 ECLYHLRGHSLRIYSVLYEPERNICISGSMDKSIRVWDLSTGTCKYVLEGHDAFVTLLNV 597

Qy     472 KDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLW 531
             | |:||:||||::|||: ||: |  |   : : |:      :    | | :||:||||
Db     598 FQNRLISGSADSTIRIWDLNTGKPLMVLPSNSGYISSFV----SDEHKIISGNDGSVKLW 653

Qy     532 DLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDF 584
           |::||: :| |:|         :| :     : | ||      ::  | |::|
Db     654 DVRTGKLLRFLLT-----DLTKIWHVDFDAMRCVAAV---QRDDQAYLEVINF 698

RESULT   5
T16607
hypothetical protein K10B2.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16607
R;Miller, N.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of C. elegans cosmid K10B2.
A;Reference number: Z18545
A;Accession: T16607
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-701 <MIL>
A;Cross-references: EMBL:U28730; NID:g860694; PID:g860695; PIDN:AAA68258.1; CESP:K10B
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:K10B2.1
A;Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3


  Query Match             20.3%;  Score 639.5;  DB 2;  Length 701;
  Best Local Similarity   33.5%;  Pred. No. 2.2e-42;
  Matches  154;  Conservative   80;  Mismatches  167;  Indels   59;  Gaps   12;

Qy     126 WSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQ 185
           ||  |:|  :|:::      |:  :   | |  |||||| ||  |   :| |    | |:
Db      83 WSEHEQLDFMDKIVHRLSHYQLGKVDNFIRPMLQRDFISNLPAHLVELIL-FNVNSDSLK 141

Qy     186 AAQ------TC-----RYWRILAEDNL----LW---------------------REKCKE 209
           : :       |     ::|: | | |:    ||                     |  |::
Db     142 SCEEVSTSWRCALARGQHWKKLIEKNVRSDSLWWGLSEKRQWDKFLNISRDMSVRRICEK 201

Qy     210 EGIDEPLHIKRRKVIKPGFIHSPWKSAY---IRQ-HRIDTNWRRGELKSPKV-LKGHDDH 264
              |  ::||| |: :   :|  :   |   ||  | || ||:||  |  ::  :  :
Db     202 FNYD--VNIKRDKLDQLILMHVFYSKLYPKIIRDIHNIDNNWKRGNYKMTRINCQSENSK 259

Qy     265 VITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRT 324
            : |||:  ::||||  |||:|:|      | | | |||| |   |  : :|||||:| |
Db     260 GVYCLQYDDDKIVSGLRDNTIKIWDRKDYSCSRILSGHTGSVLCLQYDNRVIISGSSDAT 319
```

```
Qy      325 LKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL---HVLMG 381
            ::||: |||||| ||  |   |  :      :|: |:| :: |||: : : :    ||:|
Db      320 VRVWDVETGECIKTLIHHCEAVLHLRFANGIMVTCSKDRSIAVWDMVSPRDITIRRVLVG 379

Qy      382 HVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSL 441
            | |||  | :| | :|| : |  :|||  :|   : || ||   :  ||: |  |||||
Db      380 HRAAVNVVDFDDRYIVSASGDRTIKVWSMDTLEFVRTLAGHRRGIACLQYRGRLVVSGSS 439

Qy      442 DTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQ------- 494
            | :||:||: :| |:  | ||: |   :   :  :|||  |  :|:||::
Db      440 DNTIRLWDIHSGVCLRVLEGHEELVRCIRFDEKRIVSGAYDGKIKVWDLQAALDPRALSS 499

Qy      495 --CLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWD 532
              || :|    :|   |  |||:   :::|| | |: :||
Db      500 EICLCSLV---QHTGRVFRLQFDDFQIVSSSHDDTILIWD 536

RESULT   6
B48088
beta-transducin repeat-containing protein - African clawed frog
N;Alternate names: beta-Trcp
C;Species: Xenopus laevis (African clawed frog)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jul-2000
C;Accession: B48088
R;Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.
Mol. Cell. Biol. 13, 4953-4966, 1993
A;Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase are
A;Reference number: A48088; MUID:93330289
A;Accession: B48088
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-518 <SPE>
A;Cross-references: GB:M98268; NID:g295542; PIDN:AAA02810.1; PID:g295543
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
C;Keywords: duplication
F;431-462/Domain: WD repeat homology <WD1>

  Query Match             20.2%;  Score 634.5;  DB 2;  Length 518;
  Best Local Similarity   31.1%;  Pred. No. 3.6e-42;
  Matches  161;  Conservative   90;  Mismatches  214;  Indels   53;  Gaps   12;

Qy       60 VRSFSLGKKPCKVSEYTSTTGLVPCSATPT---TFGDLRAANGQGQQ---RRRITSVQPP 113
            :  ||   :|   ||        |     |   |    : |||       ::|  |
Db        1 MEGFSCSLQPPTASEREDCNRDEPPRKIITEKNTLRQTKLANGTSSMIVPKQRKLSANYE 60

Qy      114 TGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLP----KE 169
               :  :| |: ||  :::  :: ||      |  |:   ::|  |||||: ||
Db       61 KEKELCVKYFEQWSECDQVEFVEHLISRMCHYQHGHINTYLKPMLQRDFITALPARGLDH 120

Qy      170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKE--------EGIDEPLHIKRR 221
            :|  :||:|: | |  |   |: |  :  | :||::  :          |: |     ||
Db      121 IAENILSYLDAKSLCSAELVCKEWYRVTSDGMLWKKLIERMVRTDSLWRGLAE-----RR 175

Qy      222 K------VIKPGFIHSPWKSAY------IRQ--HRIDTNWRRGELKSPKV-LKGHDDHVI 266
                     ||    :|  | |      | |    |::||| |     ::  :      :
Db      176 GWGQYLFKNKPPDGKTPPNSFYRALYPKIIQDIETIESNWRCGRHSLQRIHCRSETSKGV 235

Qy      267 TCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLK 326
             |||:   :||||  |||:|:|   | :| | |:|||| |   |  : :||:||:| |::
Db      236 YCLQYDDQKIVSGLRDNTIKIWDKNTLECKRVLMGHTGSVLCLQYDERVIITGSSDSTVR 295

Qy      327 VWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL---HVLMGHV 383
            ||:  ||| ::||  |   |  :  :   :|: |:| :: |||: :   :    ||:||
Db      296 VWDVNTGEMLNTLIHHCEAVLHLRFNNGMMVTCSKDRSIAVWDMASATDITLRRVLVGHR 355

Qy      384 AAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDT 443
            |||  | :| : :|| : |  :|||:  |   : || ||   :  ||:    ||||| |
Db      356 AAVNVVDFDDKYIVSASGDRTIKVWNTSTCEFVRTLNGHKRGIACLQYRDRLVVSGSSDN 415

Qy      444 SIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTG---------Q 494
            :||:||:| | |:  | ||: |   :   :  :|||  |  :|:||:
```

```
Db      416 TIRLWDIECGACLRVLEGHEELVRCIRFDNKRIVSGAYDGKIKVWDLVAALDPRAPAGTL 475

Qy      495 CLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWD 532
            ||:||    :|   |  |||::  :::|| | |: :||
Db      476 CLRTLV---EHSGRVFRLQFDEFQIVSSSHDDTILIWD 510

RESULT   7
T45136
WD repeat protein pop1 [imported] - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
C;Accession: T45136; T40157
R;Kominami, K.; Toda, T.
submitted to the EMBL Data Library, September 1996
A;Description: Fission yeast WD repeat protein Pop1 is involved in maintenance of plo
A;Reference number: Z22925
A;Accession: T45136
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-775 <KOM>
A;Cross-references: EMBL:Y08391; PIDN:CAA69671.1
A;Experimental source: strain h- 972
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.
submitted to the EMBL Data Library, March 1998
A;Reference number: Z21842
A;Accession: T40157
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-775 <WOO>
A;Cross-references: EMBL:AL022103; PIDN:CAA17898.1; GSPDB:GN00067; SPDB:SPBC2G2.18
A;Experimental source: strain 972h-; cosmid c2G2
C;Genetics:
A;Gene: SPBC2G2.18
A;Map position: 2
A;Note: pop1+

  Query Match             19.7%;  Score 620;  DB 2;  Length 775;
  Best Local Similarity   27.5%;  Pred. No. 8.8e-41;
  Matches  153;  Conservative  100;  Mismatches  179;  Indels  124;  Gaps   16;

Qy      109 SVQPPTGLQEW----LKMFQSWSGPEKL------LALDELIDSCEPTQVKHMMQVIEPQF 158
            |: |   |       ||: |  |  |||        |  |:  |    |::: ::: | |
Db      237 SIHPSNHLSSQKNAVLKLAQLISSFEKLPESVRQYLLFHLLSRCGKHAVQNIHKILLPIF 296

Qy      159 QRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYW-RILAEDNLLW-------------- 203
            |::|::  | |:   ||: |:   |   :|   :| :::: :  ||
Db      297 QKNFLTGFPAEITNLVLTHLDAPSLCAVSQVSHHWYKLVSSNEELWKSLFLKDGFFWDSI 356

Qy      204 ----REKCKEEGIDEPLHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLK 259
                |  | |: :     :||                | |   :   |    | :|: :|
Db      357 DSKIRTMCLEQSLSACAIMKR---------------VYFRHFNLRERW----LHAPEKIK 397

Qy      260 G-----HDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDN 314
                  |   :|| |||  ::|:  :    : ::   ||  :|:|  | | ||: :   :
Db      398 RCSFPIHGVRLITKLQFDDDKIIVSTCSPRINIYDTKTGVLIRSLEEHEGDVWTFEYVGD 457

Qy      315 IIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHL--------------HEKR----- 355
             :::||||||::||:  ||||    ||||||:||: :               | |
Db      458 TLVTGSTDRTVRVWDLRTGECKQVFYGHTSTIRCIKIVQGNQSTTDTDDVEKENRPASND 517

Qy      356 -------VVSGSRDATLRVWDIETGQCL----------------------------HVLM 380
                   ::| ||| |:|:| :    ||                              |
Db      518 ANSMPPYIISSSRDCTIRLWSL---PCLDDPPFVNVNENPDQNNDFTSATTNPFYIRTLR 574

Qy      381 GHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGI--HVVS 438
            ||  :|| |   |  :|| :||  ::||   |  ||| |:||  ||||:  :      :|
Db      575 GHTDSVREVACLGDLIVSASYDGTLRVWKASTGVCLHVLRGHVGRVYSVTINPSRQQCIS 634

Qy      439 GSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNA--DSTVKIWDIKTGQCL 496
               |  ||:|::|:|  : || || :| | :    ||||| :|  |:::::||: || |
Db      635 AGTDAKIRIWNLESGELLQTLHGHSNLVSQVTFNQNILVSASAPPDTSLRVWDLNTGSCR 694
```

```
Qy     497 QTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWR 556
             |: |  |       | ::: |:: |   |::|||:::|: :|:|:|        ::|:
Db     695 DILKCPLGH----IFFQHDESKVVSGS-HSTLQLWDIRSGKLVRDLLT-----DLDIIWQ 744

Qy     557 IRASNTKLVCAVGSRN 572
           :  :    | ||   |
Db     745 VAYNENVCVAAVLRNN 760


RESULT   8
S56245
cell division control protein CDC4 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YFL009w
C;Species: Saccharomyces cerevisiae
C;Date: 02-Sep-1995 #sequence_revision 12-Apr-1996 #text_change 26-May-2000
C;Accession: S56245; S48310; A26867; S62304
R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanu
submitted to the EMBL Data Library, May 1995
A;Description: Analysis of the nucleotide sequence of chromosome VI from Saccaromyces ce
A;Reference number: S56186
A;Accession: S56245
A;Molecule type: DNA
A;Residues: 1-779 <MUR>
A;Cross-references: EMBL:D50617; NID:g836685; PIDN:BAA09229.1; PID:d1009870; PID:g836745
R;Churcher, C.
submitted to the EMBL Data Library, September 1994
A;Reference number: S48310
A;Accession: S48310
A;Molecule type: DNA
A;Residues: 1-579 <CHU>
A;Cross-references: EMBL:Z46255; NID:g559925; PIDN:CAA86341.1; PID:g559926; MIPS:YFL009w
R;Yochem, J.; Byers, B.
J. Mol. Biol. 195, 233-245, 1987
A;Title: Structural comparison of the yeast cell division cycle gene CDC4 and a related
A;Reference number: A26867; MUID:88011240
A;Accession: A26867
A;Molecule type: DNA
A;Residues: 1-459,'E',461-779 <YOC>
A;Cross-references: EMBL:X05625; NID:g3502; PIDN:CAA29113.1; PID:g3503
R;Murakami, Y.
submitted to the EMBL Data Library, April 1994
A;Reference number: S62302
A;Accession: S62304
A;Molecule type: DNA
A;Residues: 1-779 <MUW>
A;Cross-references: EMBL:D31600; NID:g836814; PIDN:BAA06495.1; PID:d1007066; PID:g836815
C;Genetics:
A;Gene: SGD:CDC4
A;Cross-references: SGD:S0001885; MIPS:YFL009w
A;Map position: 6L
C;Function:
A;Description: initiation of DNA replication; separation of the spindle pole bodies to f
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
C;Keywords: cell cycle control
F;459-494/Domain: WD repeat homology <WD1>
F;628-659/Domain: WD repeat homology <WD2>


  Query Match             19.7%;  Score 618.5;  DB 2;  Length 779;
  Best Local Similarity   29.6%;  Pred. No. 1.2e-40;
  Matches  152;  Conservative  105;  Mismatches  176;  Indels   81;  Gaps   16;

Qy     135 LDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYW- 193
           |  |: : : :::  :  :|:   :|| |: || |::| : ::|: :|:: :    : |
Db     247 LFRLVANMDRSELSDLGTLIKDNLKRDLITSLPFEISLKIFNYLQFEDIINSLGVSQNWN 306

Qy     194 RILAEDNLLWREKCKEEGIDEP-----LHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWR 248
           :|: :   ||::    |    |     |::|  :           : :::    |  ||
Db     307 KIIRKSTSLWKKLLISENFVSPKGFNSLNLKLSQKYPKLSQQDRLRLSFLENIFILKNW- 365

Qy     249 RGELKSPK------VLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGH 302
                :||       |:||   |||||||  | :::|:||  ::|: ::  | |  | ||
```

```
Db     366 ----YNPKFVPQRTTLRGHMTSVITCLQFEDNYVITGADDKMIRVYDSINKKFLLQLSGH 421

Qy     303 TGGVWS-SQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHE----KRVV 357
            ||||:       |::|||||||::||: : | | |   || |||||: : |    | :|
Db     422 DGGVWALKYAHGGILVSGSTDRTVRVWDIKKGCCTHVFKGHNSTVRCLDIVEYKNIKYIV 481

Qy     358 SGSRDATLRVWDI-----------------------ETGQCLHVLMGHVAAVRCVQYDGR 394
           :|||| || || :                       |    : || ||:|:|| |   |
Db     482 TGSRDNTLHVWKLPKESSVPDHGEEHDYPLVFHTPEENPYFVGVLRGHMASVRTVSGHGN 541

Qy     395 RVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIH----VVSGSLDTSIRVWDV 450
            ||||:||  : |||     ||: | |||:|:||  :|  |     :| |:||:||:||:
Db     542 IVVSGSYDNTLIVWDVAQMKCLYILSGHTDRIYSTIYD--HERKRCISASMDTTIRIWDL 599

Qy     451 E----TGNC----------------IHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDI 490
           |     | |                ::|| || :|   : | |  |||  || ::: ||
Db     600 ENIWNNGECSYATNSASPCAKILGAMYTLQGHTALVGLLRLSDKFLVSAAADGSIRGWD- 658

Qy     491 KTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGS 550
                 :     : : ||:|    : | ::: |::    :::|::|: :   : ::
Db     659 -ANDYSRKFSYHHTNLSAITTFYVSDNILVSGSEN-QFNIYNLRSGKLVHANILKDA--- 713

Qy     551 GGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDF 584
              :| :      || ||  ::|  :: | :|||
Db     714 -DQIWSVNFKGKTLVAAV-EKDG--QSFLEILDF 743


RESULT   9
T18521
beta transducin-like protein - Podospora anserina
C;Species: Podospora anserina
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T18521
R;Saupe, S.; Turcq, B.; Begueret, J.
Gene 162, 135-139, 1995
A;Title: A gene responsible for vegetative incompatibility in the fungus Podospora an
A;Reference number: Z18944; MUID:96009891
A;Accession: T18521
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1356 <SAU>
A;Cross-references: EMBL:L28125; NID:g607002; PID:g607003; PIDN:AAA85775.1
C;Genetics:
A;Gene: het-e1
A;Introns: 761/3


  Query Match             18.5%;  Score 582.5;  DB 2;  Length 1356;
  Best Local Similarity   28.9%;  Pred. No. 1.8e-37;
  Matches  169;  Conservative   85;  Mismatches  193;  Indels  137;  Gaps   23;

Qy     113 PTGLQEWLKM-------------FQSWSGPEKLL----------ALDELIDSCEPTQVKH 149
           |  |:| :|:             |   |  : ||          |  |  :   || ::
Db     590 PESLEEIVKLCGSFLIIRERTVYFVHQSAKDFLLGTASDKASNKASQEAFELVFPTGIED 649

Qy     150 M--------MQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYW-------- 193
           :        : |:  : :||   |      :  :   :|  |     :| ||
Db     650 VSYIIFWRSLNVMSQKLRRDIYCLNAPGFLIDNVRVPDPDPLATVRYSCIYWIDHLRDLV 709

Qy     194 ------------------RILAEDNLLWRE-----KCKEEGID-----EPL--HIKRRK- 222
                             | |    | | |     :   |||:     | |  |  | :
Db     710 SSTSSKWVHLLQDDGDIHRFLTTKYLYWLEALSLLRALPEGINAIRQLESLLGHTIRGRL 769

Qy     223 --VIKPGF----------------------IHSPWKSAYIRQHR------------IDTN 246
             ::: |:                      : :|  |   :  :            ::
Db     770 IAIVRDGYRFALSYRMIIEKAPLQAYTSALVFAPTDSMIKKIFKKEEPGWISTISVVEAE 829

Qy     247 WRRGELKSPKVLKGHDDHVIT-CLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGG 305
           |        : |:||   |::      | |: ||||| |:|:|   :|   :|| || |
Db     830 WN----ACTQTLEGHGSSVLSVAFSADGQRVASGSDDKTIKIWDTASGTGTQTLEGHGGS 885

Qy     306 VWS---SQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHE--KRVVSGS 360
           |||   |  |:  : ||| |:|:|:|:| :| |  || ||   |: :      :|| |||
```

```
Db    886 VWSVAFSPDRER-VASGSDDKTIKIWDAASGTCTQTLEGHGGRVQSVAFSPDGQRVASGS 944

Qy    361 RDATLRVWDIETGQCLHVLMGHVAAVRCVQY--DGRRVVSGAYDFMVKVWDPETETCLHT 418
           | |:::||  :| |   | || ::|  | :  ||:|| ||: |  :|:||  : ||  |
Db    945 DDHTIKIWDAASGTCTQTLEGHGSSVLSVAFSPDGQRVASGSGDKTIKIWDTASGTCTQT 1004

Qy    419 LQGHTNRVYSLQF--DGIHVVSGSLDTSIRVWDVETGNCIHTLTGH----QSLTSGMELK 472
          |:||   |:|: |  ||  | ||| | :|::||  :| |  || ||    ||:  :
Db   1005 LEGHGGSVWSVAFSPDGQRVASGSDDKTIKIWDTASGTCTQTLEGHGGWVQSVVFSPD-- 1062

Qy    473 DNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVKL 530
             : ||: | |:||||  :| | |||:|   |  :|  : |: :   | : | |||:|:
Db   1063 GQRVASGSDDHTIKIWDAVSGTCTQTLEG---HGDSVWSVAFSPDGQRVASGSIDGTIKI 1119

Qy    531 WDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGT 574
          ||  :|   :   |||  | || |  :  |      | || :||
Db   1120 WDAASGTCTQ---TLE--GHGGWVHSVAFSPDGQRVASGSIDGT 1158


RESULT  10
T46660
sulfur controller-2 protein [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000
C;Accession: T46660
R;Kumar, A.; Paietta, J.V.
Proc. Natl. Acad. Sci. U.S.A. 92, 3343-3347, 1995
A;Title: The sulfur controller-2 negative regulatory gene of Neurospora crassa encodes a
A;Reference number: Z23121; MUID:95241499
A;Accession: T46660
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-650 <KUM>
A;Cross-references: EMBL:U17251; NID:g806757; PIDN:AAA68968.1; PID:g806758
C;Genetics:
A;Gene: scon-2
A;Map position: 3
A;Introns: 75/3; 319/1; 354/1
C;Function:
A;Description: negatively regulates sulfur structural gene expression
A;Note: scon-2+ expression is dependent on CYS3 function and the binding of CYS3 to the
C;Superfamily: unassigned WD repeat proteins; WD repeat homology


  Query Match             18.3%;  Score 576.5;  DB 2;  Length 650;
  Best Local Similarity   26.3%;  Pred. No. 1.9e-37;
  Matches  155;  Conservative   84;  Mismatches  177;  Indels  173;  Gaps   17;

Qy    106 RITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDS-CEPTQVKHMMQVIEPQFQRDFIS 164
          ::||         |  :| :     : | |  ::   | | |:  : : :    : ||||
Db     71 KLTSADQQAVTHVW-SLFSAAPARHRDLMLQGILSQLCFP-QLSFVSREVNEALKIDFIS 128

Qy    165 LLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLW--------REKCKEEGIDEPL 216
           || |||  || :|:   | :|||  : || ||: : :|          || : |   ||
Db    129 ALPVELAQKVLCYLDTVSLTKAAQVSQRWRTLADSDAVWVRMCEQHVNRKCTKCGWGLPL 188

Qy    217 --------HIKRRKVIKPG----------------------------------------- 227
                  : ::|:: | |
Db    189 LERKKLRNYTRQRQLAKGGPQGRVTELADSHDSQDRSVNQHGKRPAAEAEEEDPIKKRQC 248

Qy    228 --------FIHSP----WKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNR 275
                   :  |    ||: |  : ::  ||:    |   |||||:: | ||||   |
Db    249 MAAAEASKAVTQPKTRSWKAVYRDRWQVSYNWKNSRYKL-SVLKGHENGV-TCLQLDDNI 306

Qy    276 IVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGEC 335
          : :|| | |:|:|:  | :|:||||||| |: : |  |: :|||| | |:||||  ||||
Db    307 LATGSYDTTIKIWNIETEECIRTLVGHTAGIRALQFDDSKLISGSLDHTIKVWNWHTGEC 366

Qy    336 IHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYD--G 393
          : |   || :|  :|     : ||| | |::::|  : :  : | ||   |     |
Db    367 LSTFAAHTDSVISVHFDGHLLASGSSDKTVKIFDFNSKE-TYCLKGHSDWVNSTHVDIKS 425

Qy    394 RRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---------------------- 431
```

```
          | | | : |  :|:|| :|   : | :||   |  :
Db    426 RTVFSASDDTTIKLWDLDTRQVIRTYEGHVGHVQQVLILPPEYEPDEEVLNGASQDNQDA 485

Qy    432 -----------------------DGIH--------------------------------- 435
                                  | |
Db    486 MSVSSGGSGSPSMSHAQIERAGSPGSHSSSHNLLPSSLPSGDEDVRHLYGSAFVADESRP 545

Qy    436 -----VVSGSLDTSIRVWDVETGNCIHTLTGH----QSLTSGMELKDNI-LVSGNADSTV 485
                ::| ||:::|:||  || |: || ||     ||       | | ::||  |  |
Db    546 LPPRYFMTGGLDSTMRLWDSATGRCLRTLFGHLEGVWSLAG-----DTIRVISGANDGMV 600

Qy    486 KIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLK 534
          | |: ::|:| |   |   |   |||:  : : : : |:|||::|   |
Db    601 KTWEPRSGKCDATYTG---HCGPVTCVGLSDSLMASGSEDGTIRLHSFK 646


RESULT  11
T50211
WD-repeat protein [imported] - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 02-Sep-2000
C;Accession: T50211
R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.;
submitted to the EMBL Data Library, January 2000
A;Reference number: Z25046
A;Accession: T50211
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-506 <MCD>
A;Cross-references: EMBL:AL136538; PIDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05
A;Experimental source: strain 972h(-); cosmid c30
C;Genetics:
A;Gene: SPAC29E6.01; SPDB:SPAC30.05
A;Map position: 1
A;Introns: 43/1; 74/3
C;Superfamily: unassigned WD repeat proteins; WD repeat homology


  Query Match             17.0%;  Score 533;  DB 2;  Length 506;
  Best Local Similarity   25.8%;  Pred. No. 3.5e-34;
  Matches  140;  Conservative   91;  Mismatches  204;  Indels  108;  Gaps   14;

Qy     55 DHGSEVRSFSLGKKPCKVSEYT-----STTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
          |:|  | |        |||: |     ||:  |||             |    :  ||
Db      5 DNGKNVVS--------KVSDLTSCSDFSTSSPVPC------------LNPLSHENNRIDL 44

Qy    110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
          ::                         :|: |     |  :   :      ||  : |:|
Db     45 IR-------------------------DLLASLSKEGVVAVYNHVRSLLFTDFTEVFPEE 79

Qy    170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEG--IDEPL----------- 216
          ::| | |:|:  || :     : |: | ||  :|:    ::|  ::| :
Db     80 VSLRVFSYLDQLDLCKCKLMSKRWKRLLEDPGIWKALYMQKGWFVNENVLNEFEAWRRTH 139

Qy    217 ---------HIKRRKVIKP-GFIHSP--------------WKSAYIRQHRIDTNWRRGE- 251
                    :|:: :| | | :  |              |   |     :|:||| |
Db    140 KFPQPRFENFLKQQNIIGPYGTMFLPQQFIFDSNGRPLLNWSYLYKEHAHLDSNWRHGRF 199

Qy    252 ----LKSPKVLKGHDDHV------ITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVG 301
                :| :    |         : |:|:    :|||| | |: ||   :   |  | |
Db    200 LVSTFNNPSIRFPADQDFRATLDSVYCVQYDDEIMVSGSKDRTVSVWDVNSRFILYKLYG 259

Qy    302 HTGGVWSSQM--RDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSG 359
          |:| |       | |:::|||:| |: :|: :    :   :|||  |  : : |  ::|
Db    260 HSGSVLCLDFCRRRNLLVSGSSDSTIIIWDWQNRRPLKVYFGHTDNVLGVVVSENYIISS 319

Qy    360 SRDATLRVWDIETGQ----CLHVLMGHVAAVRCVQYDGRR--VVSGAYDFMVKVWDPETE 413
          ||| | ||| ::       |:||| ||:|:|  |||  :   :|: : |  :: ||  |
Db    320 SRDHTARVWRLDATSPAEACMHVLRGHLASVNSVQYSSKTGLIVTASSDRTLRTWDITTG 379

Qy    414 TCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
           |:  :  |   :   |::|  :|||| | :||:::  :|  :  | ||: |   :   |
Db    380 HCIRIIHAHQRGIACAQYNGKFIVSGSSDLTIRIFEASSGKLLRMLQGHEDLIRTVRFND 439
```

```
Qy    474 NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDL 533
            :|||  | ||:||:  ||:    |   |   | |  |||:   :|  :    : :|:
Db    440 EKIVSGGYDGTVRIWNFNTGEQHCVLH--NSRNSRVFGLQFDHRRIIACTHSSEILVWNF 497

Qy    534 KTG 536
            |
Db    498 DDG 500


RESULT  12
A55532
myosin-heavy-chain kinase (EC 2.7.1.129) A - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-May-2000
C;Accession: A55532
R;Futey, L.M.; Medley, Q.G.; Cote, G.P.; Egelhoff, T.T.
J. Biol. Chem. 270, 523-529, 1995
A;Title: Structural analysis of myosin heavy chain kinase A from Dictyostelium. Evidence
 the beta-subunit of heterotrimeric G proteins.
A;Reference number: A55532; MUID:95122486
A;Accession: A55532
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1146 <FUT>
A;Cross-references: GB:U16856; NID:g608519; PIDN:AAA66070.1; PID:g608520
C;Genetics:
A;Introns: #status absent
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
C;Keywords: autophosphorylation; coiled coil; multimer; phosphoprotein; phosphotransfera
F;865-898/Domain: WD repeat homology <WD1>
F;950-981/Domain: WD repeat homology <WD2>
F;991-1022/Domain: WD repeat homology <WD3>
F;1031-1062/Domain: WD repeat homology <WD4>
F;1071-1102/Domain: WD repeat homology <WD5>
F;1112-1143/Domain: WD repeat homology <WD6>


  Query Match             14.2%;  Score 446.5;  DB 2;  Length 1146;
  Best Local Similarity   33.8%;  Pred. No. 7.6e-27;
  Matches   97;  Conservative   59;  Mismatches  124;  Indels    7;  Gaps    5;

Qy    251 ELKSPKVLKGHDDHVITCLQFCGNRIV-SGSDDNTLKVWSAVTG-KCLRTLVGHTGGVWS 308
          :||    ::   : | :   |   ::: :|  | | :|:      ||| |: ||   : |
Db    858 DLKCVSTIQSFRERVNSIAFFDNQKLLCAGYGDGTYRVFDVNDNWKCLYTVNGHRKSIES 917

Qy    309 SQMRDNIIISGSTDRTLKVWNAETG--ECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLR 366
               | | : | | |:||    :|  :|| || |||  | |:  :|| : | | | |::
Db    918 IACNSNYIFTSSPDNTIKVHIIRSGNTKCIETLVGHTGEVNCVVANEKYLFSCSYDKTIK 977

Qy    367 VWDIETGQCLHVLMG-HVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNR 425
          |||: | : :    | |   :: :   || : ||  | :: ||| || : |  :||| :
Db    978 VWDLSTFKEIKSFEGVHTKYIKTLALSGRYLFSGGNDQIIYVWDTETLSMLFNMQGHEDW 1037

Qy    426 VYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTV 485
          | ||     :: | | |  |::||:   :|| || || :  |   :||  | ||: |:::
Db   1038 VLSLHCTASYLFSTSKDNVIKIWDLSNFSCIDTLKGHWNSVSSCVVKDRYLYSGSEDNSI 1097

Qy    486 KIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWD 532
          |:||: | :|: |:  |  |   | ||    | :|::: ||::|:|:
Db   1098 KVWDLDTLECVYTI--PKSHSLGVKCLMVFNNQIISAAFDGSIKVWE 1142


RESULT  13
S48052
platelet-activating factor acetylhydrolase 45K chain - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 26-May-2000
C;Accession: S48052; S51606; S48075
R;Hattori, M.; Adachi, H.; Tsujimoto, M.; Arai, H.; Inoue, K.
Nature 370, 216-218, 1994
A;Title: Miller-Dieker lissencephaly gene encodes a subunit of brain platelet-activating
A;Reference number: S48052; MUID:94301405
A;Accession: S48052
A;Molecule type: mRNA
A;Residues: 1-410 <HAT>
A;Cross-references: EMBL:D30615
A;Note: the title has been revised in S48075
R;Hattori, M.; Adachi, H.; Tsujimoto, M.; Arai, H.; Inoue, K.
Nature 370, 391-391, 1994
A;Title: Miller-Dieker lissencephaly gene encodes a subunit of brain platelet-activat
A;Reference number: S48075
A;Contents: annotation
A;Note: this is a revision to the title from reference S48052
R;Hattori, M.
submitted to the EMBL Data Library, April 1994
A;Reference number: S51606
A;Accession: S51606
A;Molecule type: mRNA
A;Residues: 1-25,'A',27-153,'E',155-410 <HAW>
A;Cross-references: EMBL:D30615; NID:g516665; PIDN:BAA06305.1; PID:d1006869; PID:g551
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
F;104-137/Domain: WD repeat homology <WD1>
F;146-179/Domain: WD repeat homology <WD2>
F;188-221/Domain: WD repeat homology <WD3>
F;230-263/Domain: WD repeat homology <WD4>
F;289-325/Domain: WD repeat homology <WD5>
F;334-367/Domain: WD repeat homology <WD6>
F;376-409/Domain: WD repeat homology <WD7>


  Query Match             14.0%;  Score 439;  DB 2;  Length 410;
  Best Local Similarity   30.1%;  Pred. No. 6.9e-27;
  Matches  110;  Conservative   62;  Mismatches  133;  Indels   60;  Gaps   10;

Qy    227 GFIHSPWKSAYIRQHRI--------------DTNWRRGELKSPK----------VLKGHD 262
          | :   | |    | ::               :    |: : ||           | ||
Db     49 GLLEKKWTSVIRLQKKVMELESKLNEAKEEFTSGGPLGQKRDPKEWIPRPPEKYALSGHR 108

Qy    263 DHVITCL-QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVW--SSQMRDNIIISG 319
            |   :     : :|| |:| |:|||   ||   ||| |||  |   |      :: |
Db    109 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASC 168

Qy    320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377
          | | |:|:|: :  ||| |::||   |  : :  :   :|| ||| |:::|:::|| |:
Db    169 SADMTIKLWDFQGFECIRTMHGHDHNVSSVAIMPNGDHIVSASRDKTIKMWEVQTGYCVK 228

Qy    378 VLMGHVAAVRCVQ--YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---- 431
             ||   || |:   ||  : | : |  |:||   |: |   |: | : |  : :
Db    229 TFTGHREWVRMVRPNQDGTLIASCSNDQTVRVWVVATKECKAELREHEHVVECISWAPES 288

Qy    432 ------------------DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
                             |  ::||| | :|::||| || |: || || :   |:
Db    289 SYSSISEATGSETKKSGKPGPFLLSGSRDKTIKMWDVSTGMCLMTLVGHDNWVRGVLFHS 348

Qy    474 --NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVK 529
              ::|   | |:::|| |  :|::||   | |:  || | |:|   :|:| | | |||
Db    349 GGKFILSCADDKTLRVWDYKNKRCMKTL---NAHEHFVTSLDFHKTAPYVVTGSVDQTVK 405

Qy    530 LWDLK 534
          :|: :
Db    406 VWECR 410


RESULT  14
S36113
LIS-1 protein - human
C;Species: Homo sapiens (man)
C;Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 26-May-2000
C;Accession: S36113
R;Reiner, O.; Carrozzo, R.; Shen, Y.; Wehnert, M.; Faustinella, F.; Dobyns, W.B.; Cas
Nature 364, 717-721, 1993
A;Title: Isolation of a Miller-Dieker lissencephaly gene containing G protein beta-su
A;Reference number: S36113; MUID:93361119
A;Accession: S36113
A;Status: preliminary
```

```
A;Molecule type: mRNA
A;Residues: 1-409 <RAS>
C;Genetics:
A;Gene: LIS-1
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
F;103-136/Domain: WD repeat homology <WD1>
F;145-178/Domain: WD repeat homology <WD2>
F;187-220/Domain: WD repeat homology <WD3>
F;229-262/Domain: WD repeat homology <WD4>
F;333-366/Domain: WD repeat homology <WD5>
F;375-408/Domain: WD repeat homology <WD6>

  Query Match             13.9%;  Score 436.5;  DB 2;  Length 409;
  Best Local Similarity   34.0%;  Pred. No. 1.1e-26;
  Matches   99;  Conservative   54;  Mismatches  103;  Indels   35;  Gaps    7;

Qy      276 IVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVW--SSQMRDNIIISGSTDRTLKVWNAETG 333
            :|| |:| |:|||   ||   ||| |||  |   |      :: | | | |:|:|: :
Db      122 MVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASCSADMTIKLWDFQGF 181

Qy      334 ECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQ- 390
            ||| |::||   |  : :  :   :|| ||| |:::|:::|| |:    ||   || |:
Db      182 ECIRTMHGHDHNVSSVAIMPNGDHIVSASRDKTIKMWEVQTGYCVKTFTGHREWVRMVRP 241

Qy      391 -YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF----------------- 431
              ||  : | : |  |:||   |: |   |: | : |   ::
Db      242 NQDGTLIASCSNDQTVRVWVVATKECKAELREHEHVVECISWAPESSYSSISEATGSETK 301

Qy      432 ----DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD--NILVSGNADSTV 485
                 |  ::||| | :|::||| || |: || || :   |:        ::|   | |:
Db      302 KSGKPGPFLLSGSRDKTIKMWDVSTGMCLMTLVGHDNWVRGVLFHSGGKFILSCADDKTL 361

Qy      486 KIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVKLWDLK 534
            ::|| |  :|::||   | |:  || | |:|   :|:| | | |||:|: :
Db      362 RVWDYKNKRCMKTL---NAHEHFVTSLDFHKTAPYVVTGSVDQTVKVWECR 409


RESULT  15
T46032
WD-40 repeat regulatory protein tup1 homolog - Arabidopsis thaliana
N;Alternate names: protein T16K5.10
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 26-May-2000
C;Accession: T46032
R;Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23015
A;Accession: T46032
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-317 <RIE>
A;Cross-references: EMBL:AL132965
A;Experimental source: cultivar Columbia; BAC clone T16K5
C;Genetics:
A;Map position: 3
A;Introns: 222/3
A;Note: T16K5.10
C;Superfamily: unassigned WD repeat proteins; WD repeat homology


  Query Match             13.1%;  Score 410.5;  DB 2;  Length 317;
  Best Local Similarity   31.9%;  Pred. No. 8.4e-25;
  Matches   98;  Conservative   66;  Mismatches  106;  Indels   37;  Gaps   12;

Qy      256 KVLKGHDDHVITCLQFC--GNRIVSGSDDNTLKVWSAVT-----GKCLRTLVGHTGGV-- 306
            : |  | :  :: ::|   |  : | | | |:: ::  |      : ::   ||  |:
Db       18 QTLTSH-NRAVSSVKFSSDGRLLASASADKTIRTYTINTINDPIAEPVQEFTGHENGISD 76

Qy      307 --WSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLH--EKRVVSGSRD 362
              :||  |   |:| | |:|||:|: |||  | || |||:   |:: :     :|||| |
Db       77 VAFSSDAR--FIVSASDDKTLKLWDVETGSLIKTLIGHTNYAFCVNFNPQSNMIVSGSFD 134
```

```
Qy      363 ATLRVWDIETGQCLHVLMGHVAAVRCVQY--DGRRVVSGAYDFMVKVWDPETETCLHTLQ 420
             |:|:||: ||:|| ||  |   |  | :  ||  :|| :|| : ::||  |  |: ||
Db      135 ETVRIWDVTTGKCLKVLPAHSDPVTAVDFNRDGSLIVSSSYDGLCRIWDSGTGHCVKTLI 194

Qy      421 GHTNRVYS-LQF--DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQ---------SLTS 467
               |   | ::|  :|  :: |:|| ::|:|:: :   : | |||           |:|:
Db      195 DDENPPVSFVRFSPNGKFILVGTLDNTLRLWNISSAKFLKTYTGHVNAQYCISSAFSVTN 254

Qy      468 GMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGT 527
            |       :|||: |: | :|:: : : || |:|  :    | |    :| : : | | |
Db      255 GKR-----IVSGSEDNCVHMWELNSKKLLQKLEGHTETVMNVAC-HPTENLIASGSLDKT 308

Qy      528 VKLWDLK 534
            |::|  |
Db      309 VRIWTQK 315


Search completed: September 25, 2001, 14:42:04
Job time: 149 sec
```

THIS PAGE BLANK (USPTO)

```
                    GenCore version 4.5
            Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          September 25, 2001, 14:41:35 ; Search time 10.7 Seconds
                                                (without alignments)
                                                1885.652 Million cell updates/sec

Title:           US-09-328-877A-8
Perfect score:   3143
Sequence:        1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                               SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB ID                   Description
------------------------------------------------------------------------------
     1   1238    39.4    579  1 SE10_CAEEL           Q93794 caenorhabdi
     2  725.5    23.1    684  1 CC4_CANAL            P53699 candida alb
     3    680    21.6    605  1 POF1_SCHPO           P87053 schizosacch
     4    650    20.7    640  1 MT30_YEAST           P39014 saccharomyc
     5    645    20.5    678  1 SCOB_EMENI           Q00659 emericella
     6  641.5    20.4    542  1 FW1B_HUMAN           Q9ukb1 homo sapien
     7    641    20.4    703  1 POP2_SCHPO           O14170 schizosacch
     8  639.5    20.3    701  1 YSS1_CAEEL           Q09990 caenorhabdi
     9    638    20.3    605  1 FW1A_HUMAN           Q9y297 homo sapien
    10  634.5    20.2    518  1 TRCB_XENLA           Q91854 xenopus lae
    11    620    19.7    775  1 POP1_SCHPO           P87060 schizosacch
    12  618.5    19.7    779  1 CC4_YEAST            P07834 saccharomyc
    13  582.5    18.5   1356  1 HET1_PODAN           Q00808 podospora a
    14  576.5    18.3    650  1 SCO2_NEUCR           Q01277 neurospora
    15    516    16.4    732  1 KMHB_DICDI           P90648 dictyosteli
    16    461    14.7    362  1 WDR5_HUMAN           Q9ugp9 homo sapien
    17    460    14.6    361  1 WDS_DROME            Q9v3j8 drosophila
    18  446.5    14.2   1146  1 KMHA_DICDI           P42527 dictyosteli
    19    439    14.0    409  1 LIS1_HUMAN           P43034 homo sapien
    20    439    14.0    409  1 LIS1_MOUSE           P43035 mus musculu
    21    438    13.9    409  1 LIS1_BOVIN           P43033 bos taurus
    22    403    12.8    267  1 YAF1_SCHPO           Q09855 schizosacch
    23    396    12.6    376  1 YKY4_CAEEL           Q17963 caenorhabdi
    24  392.5    12.5    422  1 FBW2_HUMAN           Q9ukt8 homo sapien
    25    383    12.2    742  1 PKWA_THECU           P49695 thermomonos
    26  376.5    12.0    422  1 FBW2_MOUSE           Q60584 mus musculu
    27    376    12.0    501  1 YH92_CAEEL           Q23256 caenorhabdi
    28    365    11.6    515  1 YCW2_YEAST           P25382 saccharomyc
    29    361    11.5    395  1 YZLL_CAEEL           Q93847 caenorhabdi
    30    360    11.5   1693  1 Y163_SYNY3           Q55563 synechocyst
    31    355    11.3    444  1 NUDF_EMENI           Q00664 emericella
    32  353.5    11.2    714  1 YJL2_YEAST           P47025 saccharomyc
    33  342.5    10.9    473  1 PRL1_SCHPO           O13615 schizosacch
    34    332    10.6    704  1 T2D4_DROME           P49846 drosophila
    35    330    10.5    486  1 PRL1_ARATH           Q42384 arabidopsis
    36    330    10.5    577  1 TBL1_HUMAN           O60907 homo sapien
    37    326    10.4   2295  1 WDR9_HUMAN           Q9nsi6 homo sapien
    38  323.5    10.3    332  1 GBLP_DICDI           P46800 dictyosteli
    39  318.5    10.1    713  1 TUP1_YEAST           P16649 saccharomyc
    40    318    10.1    451  1 PRL1_YEAST           Q12417 saccharomyc
    41  317.5    10.1    800  1 T2D4_HUMAN           Q15542 homo sapien
    42  315.5    10.0    612  1 TIPD_DICDI           O15736 dictyosteli
    43  314.5    10.0    479  1 PRL2_ARATH           Q39190 arabidopsis
    44    312     9.9    659  1 YK16_YEAST           P36130 saccharomyc
    45  311.5     9.9    682  1 TUP1_KLULA           P56094 kluyveromyc


                               ALIGNMENTS


RESULT   1
SE10_CAEEL
ID   SE10_CAEEL     STANDARD;      PRT;   579 AA.
AC   Q93794;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   SEL-10 PROTEIN.
GN   SEL-10 OR F55B12.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Sims M.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   REVISIONS.
RC   STRAIN=BRISTOL N2;
RA   Jones S.J.M.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z79757; CAB02129.1; -.
DR   WormPep; F55B12.3; CE16120.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 5.
DR   PROSITE; PS50082; WD_REPEATS_2; 7.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   DOMAIN      113    159       F-BOX.
FT   REPEAT      245    274       WD 1.
FT   REPEAT      286    316       WD 2.
FT   REPEAT      328    356       WD 3.
FT   REPEAT      368    396       WD 4.
FT   REPEAT      408    438       WD 5.
FT   REPEAT      453    481       WD 6.
FT   REPEAT      493    522       WD 7.
SQ   SEQUENCE   579 AA;  64275 MW;  0647245AF5964663 CRC64;
```

```
  Query Match             39.4%;  Score 1238;  DB 1;  Length 579;
  Best Local Similarity   45.6%;  Pred. No. 9.4e-90;
  Matches  262;  Conservative   85;  Mismatches  162;  Indels   66;  Gaps   11;

Qy     55 DHGSEVRSFSLGKKPC----KVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSV 110
          |:| |  |:| |        |:|        :  ||:|:   ||
Db     23 DNGEE-SSYSNGSSSSYNADKLSSSRPLQHKLDLSASPSRNNDLNPR------------- 68

Qy    111 QPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKEL 170
               ::  : |:  |  |:: |   |:     | :: :  :||| |||||:| || ||
Db     69 -----VEHLIALFKDLSSAEQMDAFTRLLQESNMTNIRQLRAIIEPHFQRDFLSCLPVEL 123

Qy    171 ALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEE-------------------- 210
           : :|  |   |||: ||  : |::::| : :|:    ||
Db    124 GMKILHNLTGYDLLKVAQVSKNWKLISEIDKIWKSLGVEEFKHHPDPTDRVTGAWQGTAI 183

Qy    211 --GIDEPLHI--------KRRKVIKPG--FIHSPWKSAYIRQHRIDTNWRRGELKSPKVL 258
            |:  | ||        :  |: | |  |  :  || |:|  :|: ||    :    ||
Db    184 AAGVTIPDHIQPCDLNVHRFLKLQKFGDIFERAADKSRYLRADKIEKNWNANPIMGSAVL 243

Qy    259 KGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRD--NII 316
          :||:||||||:|   : :|:||||||||||    |: : ||||||||||:||:      |
Db    244 RGHEDHVITCMQIHDDVLVTGSDDNTLKVWCIDKGEVMYTLVGHTGGVWTSQISQCGRYI 303

Qy    317 ISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL 376
          :|||||||:|||:   |  :||| ||||||||| :    :|:|||| ||||||:|:|: |
Db    304 VSGSTDRTVKVWSTVDGSLLHTLQGHTSTVRCMAMAGSILVTGSRDTTLRVWDVESGRHL 363

Qy    377 HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIH- 435
            | || |||||||:||  |||| ||| ||:|:  |  |: || || |||||| |:
Db    364 ATLHGHHAAVRCVQFDGTTVVSGGYDFTVKIWNAHTGRCIRTLTGHNNRVYSLLFESERS 423

Qy    436 -VVSGSLDTSIRVWDV---ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIK 491
           | ||||||||||||    |   |:  | || ||||||:|: ||||| |||| |::|||
Db    424 IVCSGSLDTSIRVWDFTRPEGQECVALLQGHTSLTSGMQLRGNILVSCNADSHVRVWDIH 483

Qy    492 TGQCLQTLQGPNKHQSAVTCLQ-FNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGS 550
           | |:  | |   |:||:| || | :| | ||||||||||||:: |  ||:||||:|||:
Db    484 EGTCVHMLSG---HRSAITSLQWFGRNMVATSSDDGTVKLWDIERGALIRDLVTLDSGGN 540

Qy    551 GGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFD 585
          || :||: :::| | ||||||| |||||:::||||
Db    541 GGCIWRLCSTSTMLACAVGSRNNTEETKVILLDFD 575


RESULT   2
CC4_CANAL
ID   CC4_CANAL      STANDARD;      PRT;   684 AA.
AC   P53699;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CELL DIVISION CONTROL PROTEIN 4.
GN   CDC4.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SGY126;
RA   Shieh J.C., White A.M., Rosamond J.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC       REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC       POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC       DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
CC       VARIOUS ASPECTS OF SPORULATION. REQUIRED FOR HTA1-HTB1 LOCUS
CC       TRANSCRIPTION ACTIVATION (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X96763; CAA65538.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 4.
DR   PROSITE; PS50082; WD_REPEATS_2; 4.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Cell division; Mitosis; Sporulation; Repeat; WD repeat.
FT   DOMAIN      212    258       F-BOX.
FT   REPEAT      322    351       WD 1.
FT   REPEAT      363    391       WD 2.
FT   REPEAT      403    431       WD 3.
FT   REPEAT      442    468       WD 4 (POTENTIAL).
FT   REPEAT      478    506       WD 5.
FT   REPEAT      519    549       WD 6.
FT   REPEAT      561    589       WD 7.
SQ   SEQUENCE   684 AA;  76090 MW;  3DD65DB31293B107 CRC64;


  Query Match             23.1%;  Score 725.5;  DB 1;  Length 684;
  Best Local Similarity   30.8%;  Pred. No. 2.2e-49;
  Matches  173;  Conservative  102;  Mismatches  181;  Indels  105;  Gaps   16;

Qy    102 QQRRRITSVQ-----PPTGLQEWLKMFQSWSG-----PEKLLALDELIDSCEPTQVKHMM 151
          | :  : |::      ||   | : :    ||      :  |    |  :  || :
Db    145 QGKANLDSIENVMSNEPTTQSELVDLVTKLSGFLSEANQNHLVFKLLQKTTRPT-LSTFN 203

Qy    152 QVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYW-----------RILAEDN 200
           :|    :|| :| :| |: : :||:|: | ||  ||:|  |           ::|  |
Db    204 NLINNSLKRDILSNVPFEVTMKILSYLDYKTLLSVAQVCKKWFDIINNPDTWIKLLKRDK 263

Qy    201 LLWREKCKEEGIDEPLHI-----------------KRRKVIKPGFIHSPWKSAYIRQHRI 243
          |:  :   :  :  |  :                 |:||:|      : |     : |||
Db    264 LITDDAVIKYELQYPDQLLREWSTLPEINSAQVLYKKRKIIV-----NRWMDPKFKPHRI 318

Qy    244 DTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHT 303
                        : || : |:||||    ::|:| || : ::|  ||: :: | ||
Db    319 S-------------VSGHGNKVVTCLQHDDEKVVTGVDDKCISIYSTQTGQLMKVLEGHE 365

Qy    304 GGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRV------- 356
          ||||: :   | :::||||||::||| :||:| |   |||||:||: :    |
Db    366 GGVWALKYTGNTLVTGSTDRTVRVWNMKTGQCTHIFRGHTSTIRCLDIIHPAVIGKNQDG 425

Qy    357 -----------VSGSRDATLRVWD----------IETGQ-------CLHVLMGHVAAVRC 388
                     ::||||  : ||           ||| :        : || ||  :||
Db    426 EDIVFPEYPLLITGSRDHNIHVWKLPVVDDSQDYIETFEGEFDNPYLIAVLSGHTQSVRS 485

Qy    389 VQYDGRRVVSGAYDFMVKVWD-PETETCLHTLQGHTNRVYS--LQFDGIHVVSGSLDTSI 445
          :   |  ::||:||  |:|||  :   | | |||| :||||  : |      |||:|::|
Db    486 ISGYGNIIISGSYDSTVRVWDLLDDGHCTHVLQGHLDRVYSTAIDFHSKTCFSGSMDSNI 545

Qy    446 RVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKH 505
           ||: |||     | || ||   ::| |::|||  ||:|::||| |||:    |:|   |
Db    546 NVWNFETGELKKVLVGHASLVGLLDLVDDVLVSAAADATLRIWDAKTGELRSKLKG---H 602

Qy    506 QSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLV 565
           :|:|| : :   |:: |:   :|||::: | | |:|::  :||    :|::|    : |
Db    603 GAAITCFEHDGLRVVSGSEK-MLKLWNVEKGTFARDLLSDVTGG----IWQVRFDYKRCV 657

Qy    566 CAVGS--RNGTEETKLLVLDF 584
           ||         || : :|||
Db    658 AAVQRIINEDEGETFIEILDF 678
```

```
RESULT   3
POF1_SCHPO
ID   POF1_SCHPO     STANDARD;      PRT;   605 AA.
AC   P87053;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   F-BOX/WD-REPEAT PROTEIN POF1.
GN   POF1 OR SPAC57A10.05.C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Katayama S., Kitamura K., Toda T.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Badcock K., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB032410; BAA84528.1; -.
DR   EMBL; Z94864; CAB08168.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 2.
DR   PROSITE; PS50082; WD_REPEATS_2; 7.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   DOMAIN       107    153       F-BOX.
FT   REPEAT       271    299       WD 1.
FT   REPEAT       311    339       WD 2.
FT   REPEAT       350    379       WD 3.
FT   REPEAT       390    420       WD 4.
FT   REPEAT       432    460       WD 5.
FT   REPEAT       472    500       WD 6.
FT   REPEAT       510    538       WD 7.
SQ   SEQUENCE   605 AA;  67110 MW;  7118C9379EC5C1F0 CRC64;


  Query Match             21.6%;  Score 680;  DB 1;  Length 605;
  Best Local Similarity   29.8%;  Pred. No. 7.1e-46;
  Matches  164;  Conservative   93;  Mismatches  181;  Indels  112;  Gaps   17;

Qy     51 KRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSV 110
          |:: |    | :|::       | :   :||                    : |:|: :|
Db     31 KKEGDQSISVSAFNIS------SMHNELSGL-------------------SEKSRQRVEAV 66

Qy    111 QPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKEL 170
                  |   |   |  |: |||  ::::|  : :      ::   : ||:|||| |:
Db     67 --------W-AAFSEASCSERKLALQGILNNCSSSLLSFASSTLDSLVRLDFLSLLPVEI 117

Qy    171 ALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKE-------------------- 209
```

```
          :  :||||: : | ||||  ::|: ||:|:::|   |::
Db    118 SFRILSFLDARSLCQAAQVSKHWKELADDDVIWHRMCEQHINRKCEKCGWGLPLLERNTL 177

Qy    210 ----------------EGID---EPLHIKRRKV--------------IKP---------- 226
                           |:|   |   :|: |:              :||
Db    178 YAAKASIQKRYERLTKRGVDQAHESSPVKKAKLDDYPTSSNEETISSVKPPSPNSDSKFF 237

Qy    227 -GFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTL 285
            |   |||  |  : |:: ||| |  :   || || | |: |||   | : ||| | |:
Db    238 LPFKTRPWKEVYAERCRVECNWRHGRCRQ-VVLSGHSDGVM-CLQLVRNILASGSYDATI 295

Qy    286 KVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTST 345
          ::|:  | : :  | ||: ||   |     :|||| |:|:::||  | |||  |:||| :
Db    296 RLWNLATFQQVALLEGHSSGVTCLQFDQCKLISGSMDKTIRIWNYRTSECISILHGHTDS 355

Qy    346 VRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQY--DGRRVVSGAYDF 403
          | |:      :|||| | |:::|    |: :  | ||   |  |:   |   |:||: |
Db    356 VLCLTFDSTLLVSGSADCTVKLWHFSGGKRI-TLRGHTGPVNSVRIIRDRGLVLSGSDDS 414

Qy    404 MVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQ 463
           :|:|  || |||||   |   | ||      : | ||| :|: ||:|   |:||| ||
Db    415 TIKIWSLETNTCLHTFSAHIGPVQSLALADSRLFSCSLDGTIKQWDIEKKKCVHTLFGH- 473

Qy    464 SLTSGMEL-KDNI-LVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVIT 521
           :    |:  |:: |:||  |  ||:|:    :|: ||:    |   || :      |::
Db    474 -IEGVWEIAADHLRLISGAHDGVVKVWE--ACECVHTLK---NHSEPVTSVALGDCEVVS 527

Qy    522 SSDDGTVKLW 531
           |:|| : ||
Db    528 GSEDGKIYLW 537


RESULT   4
MT30_YEAST
ID   MT30_YEAST     STANDARD;      PRT;   640 AA.
AC   P39014;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MET30 PROTEIN.
GN   MET30 OR YIL046W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X2180-1A;
RX   MEDLINE=96069360; PubMed=8524217;
RA   Thomas D., Kuras L., Barbey R., Cherest H., Blaiseau P.L.,
RA   Surdin-Kerjan Y.;
RT   "Met30p, a yeast transcriptional inhibitor that responds to S-
RT   adenosylmethionine, is an essential protein with WD40 repeats.";
RL   Mol. Cell. Biol. 15:6526-6534(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA   Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA   Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA   Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA   Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA   Walsh S.V., Whitehead S.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS
CC       GENES EXPRESSION.
CC   -!- SUBUNIT: SEEMS TO INTERACT WITH MET4.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC   -!- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z46861; CAA86905.1; -.
DR   EMBL; L26505; AAA96717.1; -.
DR   SGD; S0001308; MET30.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 4.
DR   PROSITE; PS50082; WD_REPEATS_2; 6.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Transcription regulation; Methionine biosynthesis;
KW   Cysteine biosynthesis; Repeat; WD repeat.
FT   DOMAIN      181    227       F-BOX.
FT   REPEAT      300    328       WD 1.
FT   REPEAT      340    368       WD 2.
FT   REPEAT      380    408       WD 3.
FT   REPEAT      419    449       WD 4.
FT   REPEAT      461    499       WD 5.
FT   REPEAT      509    538       WD 6.
FT   REPEAT      550    578       WD 7.
FT   REPEAT      607    635       WD 8 (POTENTIAL).
FT   CONFLICT     61     61       M -> I (IN REF. 1).
SQ   SEQUENCE   640 AA;  72835 MW;  5135D4BCA2E1EB97 CRC64;

  Query Match             20.7%;  Score 650;  DB 1;  Length 640;
  Best Local Similarity   26.4%;  Pred. No. 1.8e-43;
  Matches  163;  Conservative  114;  Mismatches  197;  Indels  144;  Gaps   18;

Qy       7 PTLNHGLVPVDLKSAKEPLPHQTVMKIFSISIIAQGLPFCRRRMKRKLDHGSEVRSFSLG 66
           |: :  |   |  :  :|||     |            || |       |  ::: ||
Db      68 PSSSPDLATNDSGTRVQPLPEYNFTK------------FCYR-------HNPDIQ-FSPT 107

Qy      67 KKPC------KVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWL 120
              |      :  |  :    :|                          :|  : :   :
Db     108 HTACYKQDLKRTQEINANIAKLP-------------------------LQEQSDIHHII 141

Qy     121 KMFQSWSGPEKLLALDELID-SCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLE 179
             : : :   : | || ::  || | |: ::  ::    : ||||:||:||:| :||:|:
Db     142 SKYSNSNDKIRKLILDGILSTSCFP-QLSYISSLVTHMIKIDFISILPQELSLKILSYLD 200

Qy     180 PKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEP----------LHIKRRKVIKPGFI 229
            : |  | : || |: ||:|: :|   | |: ||            ||:|| :: :
Db     201 CQSLCNATRVCRKWQKLADDDRVWYHMC-EQHIDRKCPNCGWGLPLLHMKRARIQQNSTG 259

Qy     230 HS-----------PWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVS 278
            |           |||  |  : ::::|||:|  :  :  ||| | |:| |||    : :
Db     260 SSSNADIQTQTTRPWKVIYRERFKVESNWRKGHCRIQE-FKGHMDGVLT-LQFNYRLLFT 317

Qy     279 GSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHT 338
           || |:|: :|   ||| :| | ||: || :    |  :|:|| |:|::|||  ||||| |
Db     318 GSYDSTIGIWDLFTGKLIRRLSGHSDGVKTLYFDDRKLITGSLDKTIRVWNYITGECIST 377

Qy     339 LYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGR--RV 396
             ||: :|  :  ::| :|||| | |::|| :|:  | : | ||   | ||:   :
Db     378 YRGHSDSVLSVDSYQKVIVSGSADKTVKVWHVESRTC-YTLRGHTEWVNCVKLHPKSFSC 436

Qy     397 VSGAYDFMVKVWDPETETCLHTLQGHTNRVYSL--------------------------- 429
            | : |  :::||  |  :||   :||  :|  :
Db     437 FSCSDDTTIRMWDIRTNSCLKVFRGHVGQVQKIIPLTIKDVENLATDNTSDGSSPQDDPT 496

Qy     430 QFDG----------------------IHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTS 467
             ||                       |::|  || :|::|||:|| || |  ||
Db     497 MTDGADESDTPSNEQETVLDENIPYPTHLLSCGLDNTIKLWDVKTGKCIRTQFGHVEGVW 556
```

```
Qy     468 GMELKDNILVSGNADSTVKIWDIKTGQCLQT-----LQGPNKHQ---------SAVTCLQ 513
            :   :  ::||: | ::|:||:::|:|: |     ||   :|          : : |:
Db     557 DIAADNFRIISGSHDGSIKVWDLQSGKCMHTFNGRRLQRETQHTQTQSLGDKVAPIACVC 616

Qy     514 FNKNFVITSSDDGTVKLW 531
              :   :  : | ||::
Db     617 IGDSECFSGDEFGCVKMY 634


RESULT   5
SCOB_EMENI
ID   SCOB_EMENI     STANDARD;      PRT;   678 AA.
AC   Q00659;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   SULFUR METABOLITE REPRESSION CONTROL PROTEIN.
GN   SCONB OR MAPB1.
OS   Emericella nidulans (Aspergillus nidulans).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=5072;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Natorff R.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE
CC       REPRESSION.
CC   -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC       PROTEINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U21220; AAC15905.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 4.
DR   PROSITE; PS50082; WD_REPEATS_2; 7.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Transcription regulation; Repeat; WD repeat.
FT   DOMAIN      178    224       F-BOX.
FT   REPEAT      347    375       WD 1.
FT   REPEAT      387    415       WD 2.
FT   REPEAT      427    455       WD 3.
FT   REPEAT      466    496       WD 4.
FT   REPEAT      508    543       WD 5.
FT   REPEAT      553    595       WD 6.
FT   REPEAT      607    635       WD 7.
FT   REPEAT      647    675       WD 8.
SQ   SEQUENCE   678 AA;  76070 MW;  D840D452E37B4C53 CRC64;

  Query Match             20.5%;  Score 645;  DB 1;  Length 678;
  Best Local Similarity   27.2%;  Pred. No. 4.7e-43;
  Matches  175;  Conservative   86;  Mismatches  210;  Indels  172;  Gaps   17;

Qy      55 DHGSEVRSFSLGKKPCKVSEYTSTTGL--------VPCSATPTTFGDLRAANG------Q 100
           ||     |||  | | | :       |         |  |      || :||       :
Db      46 DHN---HSFSNAKAPAKFANQNVAPFLARHIPEQYAPLGAQSILPADLSSANSKYCYRHR 102
```

```
Qy     101 GQQRRRITSVQP-------------PTGLQE-----WLKMFQSWSGPEKLLALDELIDSCE 143
             |: |  : :|            | | |:     |  :| :     : | |  ::  |
Db     103 PDQKCRRQADEPSMDKLQRELESLPQGDQQSISHVW-SLFSAAPAKHRKLILQGIMAQCC 161

Qy     144 PTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLW 203
             |: ::   :    : |||: || |:|  :| :|:   | :|:|  | ||  ||:|:::|
Db     162 FPQLSYISATVRDLIRIDFITALPPEIAFKILCYLDTTSLCKASQVSRGWRALADDDVVW 221

Qy     204 --------REKCKEEGIDEPLHIKRR---------------------------------- 221
                     |||: |   ||  ::|
Db     222 HRMCEQHIHRKCKKCGWGLPLLDRKRLRESKREIELRATTWDKGVVGPRSPDASAESPPS 281

Qy     222 ----------KVIK---------------PGFI---HSPWKSAYIRQHRIDTNWRRGELK 253
                      |:|                 |    : |||  |  : ::  |||: |
Db     282 GKRKLEDDEVAVVKRHCSSLGSDAGVDKDSDFFKTRYRPWKEVYKDRFKVGTNWKYGRC- 340

Qy     254 SPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRD 313
           | |  ||| : |: ||||  | : :|| | |:|:|   ||: |||| ||  |:   |  |
Db     341 SIKTFKGHTNGVM-CLQFEDNILATGSYDTTIKIWDTETGEELRTLRGHESGIRCLQFDD 399

Qy     314 NIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETG 373
             :|||| |||:||||  |||||  |  ||   |  :|     : |||  | |:::|: |
Db     400 TKLISGSMDRTIKVWNWRTGECISTYTGHRGGVIGLHFDASILASGSVDKTVKIWNFE-D 458

Qy     374 QCLHVLMGHVAAVRCVQYD--GRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRV----- 426
           :    | ||   |  |: |   | | | : |  |::|| :|:||: |  ||  :|
Db     459 KSTFSLRGHTDWVNAVRVDTSSRTVFSASDDCTVRLWDLDTKTCIRTFHGHVGQVQQVVP 518

Qy     427 ------------------------------------------------YSLQFDG----- 433
                                                           :   ||
Db     519 LPREFEFEEHDAECENDDLSTTSGDANPPSIQASMGLEPNAAYSQSSAFGTSFDNGRAAP 578

Qy     434 -IHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKT 492
             ::|: :||::||:|:  || |: |  ||      :      :|||  |  :|||| :|
Db     579 PRYMVTSALDSTIRLWETTTGRCLRTFFGHLEGVWALGADTLRIVSGAEDRMIKIWDPRT 638

Qy     493 GQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKT 535
           |:| :|  |   |   |||:    :   | |:|  |:::  ::
Db     639 GKCERTFTG---HSGPVTCIGLGDSRFATGSEDCEVRMYSFQS 678


RESULT   6
FW1B_HUMAN
ID   FW1B_HUMAN     STANDARD;      PRT;   542 AA.
AC   Q9UKB1; Q9Y4C6; Q9P2S8; Q9P2S9;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2).
GN   FBXW1B OR FBW1B OR BTRCP2 OR KIAA0696.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cenciarelli C., Chiaur D.S., Guardavaccaro D., Parks W., Vidal M.,
RA   Pagano M.;
RT   "Identification of a family of human F-box proteins.";
RL   Curr. Biol. 9:1177-1179(1999).
RN   [2]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Fetal lung;
RA   Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,
RA   Katoh M.;
RT   "Molecular cloning and genomic structure of the betaTRCP2 gene on
RT   chromosome 5q35.1.";
RL   Biochem. Biophys. Res. Commun. 269:103-109(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98403880; PubMed=9734811;
RA   Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
```

```
RA   Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
CC   -!- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PROTEINS AND
CC       PROMOTES THEIR UBIQUITINATION AND DEGRADATION.
CC   -!- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC   -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A, B AND C (SHOWN HERE); ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF176022; AAF04528.1; -.
DR   EMBL; AB033279; BAA92329.1; -.
DR   EMBL; AB033280; BAA92330.1; -.
DR   EMBL; AB033281; BAA92331.1; -.
DR   EMBL; AB014596; BAA31671.1; ALT_INIT.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 5.
DR   PROSITE; PS50082; WD_REPEATS_2; 7.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Ubiquitin conjugation; Repeat; WD repeat; Alternative splicing.
FT   DOMAIN      129    167       F-BOX.
FT   REPEAT      238    275       WD 1.
FT   REPEAT      278    315       WD 2.
FT   REPEAT      318    355       WD 3.
FT   REPEAT      361    398       WD 4.
FT   REPEAT      401    440       WD 5.
FT   REPEAT      442    478       WD 6.
FT   REPEAT      490    527       WD 7.
FT   VARSPLIC     16     49       MISSING (IN ISOFORM A).
FT   VARSPLIC     16     48       CSVPRSLWLGCANLVESMCALSCLQSMPSVRCL -> NTSV
FT                                MEDQNEDESPKKNTLW (IN ISOFORM B).
SQ   SEQUENCE   542 AA;  62090 MW;  7CD40087EFAA5C8A CRC64;


  Query Match             20.4%;  Score 641.5;  DB 1;  Length 542;
  Best Local Similarity   32.1%;  Pred. No. 6.5e-43;
  Matches  143;  Conservative   89;  Mismatches  177;  Indels   37;  Gaps    9;

Qy     120 LKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE----LALYVL 175
           :| |  ||  :::  :: ||      |  |:   ::|  |||||: ||::    :|  :|
Db      77 IKYFDQWSESDQVEFVEHLISRMCHYQHGHINSYLKPMLQRDFITALPEQGLDHIAENIL 136

Qy     176 SFLEPKDLLQAAQTCRYW-RILAEDNL-------------LWREKCKEEGIDEPLHIKRR 221
           |:|: : |  |   |: | |:::|  |             ||:   :  | |: |  | |
Db     137 SYLDARSLCAAELVCKEWQRVISEGMLWKKLIERMVRTDPLWKGLSERRGWDQYL-FKNR 195

Qy     222 KVIKP--GFIHSPWKSAYIRQHRIDTNWRRGELKSPKV-LKGHDDHVITCLQFCGNRIVS 278
               |   |  | :         |::||| |     ::  :  :   : |||:   :|:|
Db     196 PTDGPPNSFYRSLYPKIIQDIETIESNWRCGRHNLQRIQCRSENSKGVYCLQYDDEKIIS 255

Qy     279 GSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHT 338
           |  ||::|:|   : :||: | |||| |   |  : :|::||:| |::||:  ||| ::|
Db     256 GLRDNSIKIWDKTSLECLKVLTGHTGSVLCLQYDERVIVTGSSDSTVRVWDVNTGEVLNT 315

Qy     339 LYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL---HVLMGHVAAVRCVQYDGRR 395
           |  |   |  :      :|: |:| :: |||: :   :    ||:|| |||  | :| :
```

```
Db    316 LIHHNEAVLHLRFSNGLMVTCSKDRSIAVWDMASATDITLRRVLVGHRAAVNVVDFDDKY 375

Qy    396 VVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNC 455
          :|| : |  :|||   |   : || ||   :  ||:    ||||| | :||:||:| | |
Db    376 IVSASGDRTIKVWSTSTCEFVRTLNGHKRGIACLQYRDRLVVSGSSDNTIRLWDIECGAC 435

Qy    456 IHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTG---------QCLQTLQGPNKHQ 506
          :  | ||: |   :   :  :|||  |  :|:||::            ||:||    :|
Db    436 LRVLEGHEELVRCIRFDNKRIVSGAYDGKIKVWDLQAALDPRAPASTLCLRTLV---EHS 492

Qy    507 SAVTCLQFNKNFVITSSDDGTVKLWD 532
            |  |||::  :|:|| | |: :||
Db    493 GRVFRLQFDEFQIISSSHDDTILIWD 518

RESULT   7
POP2_SCHPO
ID   POP2_SCHPO     STANDARD;      PRT;   703 AA.
AC   O14170;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WD-REPEAT PROTEIN POP2 (PROTEOLYSIS FACTOR SUD1).
GN   POP2 OR SUD1 OR SPAC4D7.03.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Wolf D.A., Jackson P.K.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98318628; PubMed=9653157;
RA   Jallepalli P.V., Tien D., Kelly T.J.;
RT   "sud1+ targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1
RT   proteins for degradation and stops unwanted diploidization in fission
RT   yeast.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PUTATIVE COMPONENT OF E3 (PROTEIN-UBIQUITIN LIGASE)
CC       COMPLEX; TARGETS CYCLIN-DEPENDENT KINASE-PHOSPHORYLATED CDC18 AND
CC       RUM1 PROTEINS FOR DEGRADATION.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF038867; AAB95480.1; -.
DR   EMBL; AF064515; AAC39496.1; -.
DR   EMBL; Z98602; CAB11275.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 3.
DR   PROSITE; PS50082; WD_REPEATS_2; 6.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
```

```
KW   Repeat; WD repeat.
FT   DOMAIN      236    283       F-BOX.
FT   REPEAT      389    417       WD 1.
FT   REPEAT      429    473       WD 2.
FT   REPEAT      505    533       WD 3.
FT   REPEAT      545    575       WD 4.
FT   REPEAT      587    615       WD 5.
FT   REPEAT      625    654       WD 6.
SQ   SEQUENCE   703 AA;  79605 MW;  0336A0568C152396 CRC64;

  Query Match             20.4%;  Score 641;  DB 1;  Length 703;
  Best Local Similarity   27.3%;  Pred. No. 1e-42;
  Matches  178;  Conservative  111;  Mismatches  248;  Indels  116;  Gaps   17;

Qy      9 LNHGLVPVDLKSAKEPL-----------PHQTVMKIFSISIIAQ--------GLP----F 45
          | |   |  |:   | :           |  :|   ||||:: |         ||
Db     85 LKHSCTPSFLRKFNESIENVSYKCLDHSPPDSVPGDFSISLVPQRNFLYSHSSLPPKIIS 144

Qy     46 CRRRMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRR 105
            |  : |||:     | :    |   :  | : |  | |       ||   | |   :
Db    145 IDRNNRIKLDNSISSNSDNFPPSPKVDTSNTVSPGSKPISE------DLEDLNLQSIVQ- 197

Qy    106 RITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISL 165
            |    | |:| :              |  :|: ||    :: ::   ||  ::| :|
Db    198 --TFEDLPEGIQSY--------------AFFQLLRSCNRQSMRLLLNECEPLLKKDILSN 241

Qy    166 LPKELALYVLSFLEPKDLLQAAQTCRYW-RILAEDNLLWREKCKEEGI---DEPLHIKRR 221
          ||  :   :|  |:    |        | |||      |:      |    :
Db    242 LPFSIVQSILLNLDIHSFLSCRLVSPTWNRILDVHTSYWKHMFSLFGFQINENDWKYANP 301

Qy    222 KVIKPGFIHSPWKS-----AYIRQHRIDTNWRRGELKSPKVLKGH-------DDHVITCL 269
           : :| |:|:   |        ::| ::    |     | :   |        : :|| |
Db    302 NLNRPPFLHNDQISDDYFPEIFKRHFLN----RKRWLFPSIPPSHLSFPIHVPNFMITSL 357

Qy    270 QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWN 329
              :||:: |   |::: :|:||     | ||  |||: :: :| ::||| |:|::|||
Db    358 LLHKDRIITTSGSGTIQIHNAITGVLEARLEGHKEGVWAVKIHENTLVSGSIDKTVRVWN 417

Qy    330 AETGECIHTLYGHTSTVRCMHL----------------HEKRVVSGSRDATLRVWDI--- 370
           |  :| |   || | :||: :                 :  :|||||| ||||| :
Db    418 IEKAKCTHIFRGHISIIRCLEILVPSRLIRHGVEIVEPDQPYIVSGSRDHTLRVWKLPKN 477

Qy    371 -----------------ETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETE 413
                           :    :| |:||  :|| :   |  :|||:||  :::|   |
Db    478 TDPPYLPDNTNSIDRWEKNPYFVHTLIGHTDSVRTISGYGDILVSGSYDSSIRIWRVSTG 537

Qy    414 TCLHTLQGHTNRVYSLQFDGIH--VVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMEL 471
           ||: |:||: |:||: ::      :|||:| ||||||: || | : | || :  : : :
Db    538 ECLYHLRGHSLRIYSVLYEPERNICISGSMDKSIRVWDLSTGTCKYVLEGHDAFVTLLNV 597

Qy    472 KDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLW 531
            | |:||:||||::|||: ||: |  |   : : |:      :    | | :||:||||
Db    598 FQNRLISGSADSTIRIWDLNTGKPLMVLPSNSGYISSFV----SDEHKIISGNDGSVKLW 653

Qy    532 DLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDF 584
          |::||: :| |:|         :| :     : | ||      ::  | |::|
Db    654 DVRTGKLLRFLLT-----DLTKIWHVDFDAMRCVAAV---QRDDQAYLEVINF 698

RESULT   8
YSS1_CAEEL
ID   YSS1_CAEEL     STANDARD;      PRT;   701 AA.
AC   Q09990;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   HYPOTHETICAL 80.3 KDA TRP-ASP REPEATS CONTAINING PROTEIN K10B2.1 IN
DE   CHROMOSOME II.
GN   K10B2.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
```

```
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Miller N.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   -!- SIMILARITY: STRONG, TO X.LAEVIS BETA-TRCP.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U28730; AAA68258.1; -.
DR   WormPep; K10B2.1; CE02008.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 5.
DR   PROSITE; PS50082; WD_REPEATS_2; 7.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Hypothetical protein; Repeat; WD repeat.
FT   REPEAT       256    284       WD 1.
FT   REPEAT       296    324       WD 2.
FT   REPEAT       336    364       WD 3.
FT   REPEAT       379    407       WD 4.
FT   REPEAT       419    447       WD 5.
FT   REPEAT       459    487       WD 6.
FT   REPEAT       508    536       WD 7.
FT   DOMAIN       606    615       POLY-ALA.
SQ   SEQUENCE   701 AA;  80320 MW;  69FA0B00F83270E3 CRC64;


  Query Match             20.3%;  Score 639.5;  DB 1;  Length 701;
  Best Local Similarity   33.5%;  Pred. No. 1.3e-42;
  Matches  154;  Conservative   80;  Mismatches  167;  Indels   59;  Gaps   12;

Qy    126 WSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQ 185
          ||  |:|  :|:::      |:  :   | |  |||||| ||  |   :| |    | |:
Db     83 WSEHEQLDFMDKIVHRLSHYQLGKVDNFIRPMLQRDFISNLPAHLVELIL-FNVNSDSLK 141

Qy    186 AAQ------TC-----RYWRILAEDNL----LW---------------------REKCKE 209
          : :       |     ::|: | | |:    ||                     |  |::
Db    142 SCEEVSTSWRCALARGQHWKKLIEKNVRSDSLWWGLSEKRQWDKFLNISRDMSVRRICEK 201

Qy    210 EGIDEPLHIKRRKVIKPGFIHSPWKSAY---IRQ-HRIDTNWRRGELKSPKV-LKGHDDH 264
             |  ::||| |: :   :|  :   |   ||  | || ||:||  |  ::  :  :
Db    202 FNYD--VNIKRDKLDQLILMHVFYSKLYPKIIRDIHNIDNNWKRGNYKMTRINCQSENSK 259

Qy    265 VITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRT 324
           : |||:  ::||||  |||:|:|      | | | |||| |   |  : :|||||:| |
Db    260 GVYCLQYDDDKIVSGLRDNTIKIWDRKDYSCSRILSGHTGSVLCLQYDNRVIISGSSDAT 319

Qy    325 LKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL---HVLMG 381
          ::||: |||||| ||  |   |  :      :|: |:| :: |||: : : :    ||:|
Db    320 VRVWDVETGECIKTLIHHCEAVLHLRFANGIMVTCSKDRSIAVWDMVSPRDITIRRVLVG 379

Qy    382 HVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSL 441
          | |||  | :| | :|| : |  :|||  :|   : || ||   :  ||: |  |||||
Db    380 HRAAVNVVDFDDRYIVSASGDRTIKVWSMDTLEFVRTLAGHRRGIACLQYRGRLVVSGSS 439

Qy    442 DTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQ------- 494
          | :||:||: :| |:  | ||: |   :   :  :|||  |  :|:||::
Db    440 DNTIRLWDIHSGVCLRVLEGHEELVRCIRFDEKRIVSGAYDGKIKVWDLQAALDPRALSS 499

Qy    495 --CLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWD 532
            || :|    :|   |  |||:   :::|| | |: :||
Db    500 EICLCSLV---QHTGRVFRLQFDDFQIVSSSHDDTILIWD 536
```

```
RESULT   9
FW1A_HUMAN
ID   FW1A_HUMAN     STANDARD;      PRT;   605 AA.
AC   Q9Y297; Q9Y213;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP)
DE   (E3RSIKAPPAB) (PIKAPPABALPHA-E3 RECEPTOR SUBUNIT).
GN   FBXW1A OR FBW1A OR BTRCP OR BTRC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=99075339; PubMed=9859996;
RA   Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA   Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;
RT   "Identification of the receptor component of the IkappaBalpha-
RT   ubiquitin ligase.";
RL   Nature 396:590-594(1998).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Lymphoid;
RX   MEDLINE=98325370; PubMed=9660940;
RA   Margottin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,
RA   Thomas D., Strebel K., Benarous R.;
RT   "A novel human WD protein, h-beta TrCp, that interacts with HIV-1 Vpu
RT   connects CD4 to the ER degradation pathway through an F-box motif.";
RL   Mol. Cell 1:565-574(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RA   Cenciarelli C., Chiaur D.S., Guardavaccaro D., Parks W., Vidal M.,
RA   Pagano M.;
RT   "Identification of a family of human F-box proteins.";
RL   Curr. Biol. 9:1177-1179(1999).
RN   [4]
RP   CHARACTERIZATION.
RA   Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,
RA   Harper J.W.;
RT   "The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically
RT   with phosphorylated destruction motifs in I-kappa-B-alpha and
RT   beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro.";
RL   Genes Dev. 13:270-283(1999).
CC   -!- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA
CC       (PIKAPPABALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR
CC       UBIQUITINATION AND DEGRADATION.
CC   -!- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF101784; AAD08702.1; -.
DR   EMBL; Y14153; CAA74572.1; -.
DR   EMBL; AF129530; AAF04464.1; -.
DR   MIM; 603482; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 7.
```

```
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 6.
DR   PROSITE; PS50082; WD_REPEATS_2; 7.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Ubiquitin conjugation; Repeat; WD repeat; Alternative splicing.
FT   DOMAIN      190    228       F-BOX.
FT   REPEAT      301    338       WD 1.
FT   REPEAT      341    378       WD 2.
FT   REPEAT      381    418       WD 3.
FT   REPEAT      424    461       WD 4.
FT   REPEAT      464    503       WD 5.
FT   REPEAT      505    541       WD 6.
FT   REPEAT      553    590       WD 7.
FT   VARSPLIC     17     52       MISSING (IN ISOFORM 2).
SQ   SEQUENCE   605 AA;  68866 MW;  4C67F3B7E400FD37 CRC64;

  Query Match             20.3%;  Score 638;  DB 1;  Length 605;
  Best Local Similarity   30.8%;  Pred. No. 1.4e-42;
  Matches  166;  Conservative   92;  Mismatches  209;  Indels   72;  Gaps   15;

Qy        56 HGSEVRSFSLGKKPCKV--------SEYTSTTGL-----VPCSA-----TPTTFGDLRAA 97
             : ||    : |: | |:          | |   |       | |     |       : |
Db        53 NSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTENCVAKTKLA 112

Qy        98 NGQG-----QQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQ 152
             ||       :||:   | :    |   :| |: ||  :::  :: ||      |  |:
Db       113 NGTSSMIVPKQRKLSASYEKEKEL--CVKYFEQWSESDQVEFVEHLISQMCHYQHGHINS 170

Qy       153 VIEPQFQRDFISLLP----KELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCK 208
              ::|  |||||: ||      :|  :||:|: | |  |   |: |  :  | :||::  :
Db       171 YLKPMLQRDFITALPARGLDHIAENILSYLDAKSLCAAELVCKEWYRVTSDGMLWKKLIE 230

Qy       209 E--------EGIDEPLHIKRRK------VIKPGFIHSPWKSAY------IRQ--HRIDTN 246
                       |: |     ||         ||   ::|  | |      | |    |::|
Db       231 RMVRTDSLWRGLAE-----RRGWGQYLFKNKPPDGNAPPNSFYRALYPKIIQDIETIESN 285

Qy       247 WRRGELKSPKV-LKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGG 305
             || |     ::  :      : |||:   :||||  |||:|:|   | :| | | ||||
Db       286 WRCGRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECKRILTGHTGS 345

Qy       306 VWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATL 365
             |   |  : :||:||:| |::||:  ||| ::||  |   |  :  :   :|: |:| ::
Db       346 VLCLQYDERVIITGSSDSTVRVWDVNTGEMLNTLIHHCEAVLHLRFNNGMMVTCSKDRSI 405

Qy       366 RVWDIETGQCL---HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGH 422
              |||: :   :    ||:|| |||  | :| : :|| : |  :|||:  |   : || ||
Db       406 AVWDMASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRTIKVWNTSTCEFVRTLNGH 465

Qy       423 TNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNAD 482
                :  ||:    ||||| | :||:||:| | |:  | ||: |   :   :  :|||  |
Db       466 KRGIACLQYRDRLVVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDNKRIVSGAYD 525

Qy       483 STVKIWDIKTG---------QCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWD 532
               :|:||:             ||:||    :|   |  |||::  :::|| | |: :||
Db       526 GKIKVWDLVAALDPRAPAGTLCLRTLV---EHSGRVFRLQFDEFQIVSSSHDDTILIWD 581

RESULT  10
TRCB_XENLA
ID   TRCB_XENLA     STANDARD;      PRT;   518 AA.
AC   Q91854; P70037; P70038;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   BETA-TRCP (BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN).
GN   FBXW1 OR BTRCP.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93330289; PubMed=8393141;
RA   Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
RT   "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
RT   anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein
RT   with beta-transducin repeats.";
RL   Mol. Cell. Biol. 13:4953-4966(1993).
RN   [2]
RP   SEQUENCE OF 302-518 FROM N.A.
RX   MEDLINE=97109804; PubMed=8952061;
RA   Hudson J.W., Alarcon V.B., Elinson R.P.;
RT   "Identification of new localized RNAs in the Xenopus oocyte by
RT   differential display PCR.";
RL   Dev. Genet. 19:190-198(1996).
CC   -!- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PHOSPHARYLATED
CC       PROTEINS AND PROMOTES THEIR UBIQUITINATION AND DEGRADATION.
CC   -!- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC       (BY SIMILARITY).
CC   -!- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
CC       MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
CC       GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
CC       NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
CC       TADPOLE EMBRYO.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   -!- SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M98268; AAA02810.1; -.
DR   EMBL; U63921; AAB49671.1; -.
DR   EMBL; U63922; AAB49672.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 6.
DR   PROSITE; PS50082; WD_REPEATS_2; 7.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Ubiquitin conjugation; Repeat; WD repeat.
FT   DOMAIN      119    157       F-BOX.
FT   REPEAT      230    258       WD 1.
FT   REPEAT      270    298       WD 2.
FT   REPEAT      310    338       WD 3.
FT   REPEAT      353    381       WD 4.
FT   REPEAT      393    421       WD 5.
FT   REPEAT      433    461       WD 6.
FT   REPEAT      482    510       WD 7.
FT   CONFLICT    302    304       GEM -> EFR (IN REF. 2).
FT   CONFLICT    516    518       GLA -> AAH (IN REF. 2).
SQ   SEQUENCE   518 AA;  59507 MW;  2A52EC19028127F3 CRC64;

  Query Match             20.2%;  Score 634.5;  DB 1;  Length 518;
  Best Local Similarity   31.1%;  Pred. No. 2.2e-42;
  Matches  161;  Conservative   90;  Mismatches  214;  Indels   53;  Gaps   12;

Qy        60 VRSFSLGKKPCKVSEYTSTTGLVPCSATPT---TFGDLRAANGQGQQ---RRRITSVQPP 113
             :  ||   :|   ||        |     |   |    : |||       ::|  |
Db         1 MEGFSCSLQPPTASEREDCNRDEPPRKIITEKNTLRQTKLANGTSSMIVPKQRKLSANYE 60

Qy       114 TGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLP----KE 169
                :  :| |: ||  :::  :: ||      |  |:   ::|  |||||: ||
Db        61 KEKELCVKYFEQWSECDQVEFVEHLISRMCHYQHGHINTYLKPMLQRDFITALPARGLDH 120
```

```
Qy      170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKE--------EGIDEPLHIKRR 221
            :|  :||:|: | |  |   |: |  :  | :||::  :          |: |     ||
Db      121 IAENILSYLDAKSLCSAELVCKEWYRVTSDGMLWKKLIERMVRTDSLWRGLAE-----RR 175

Qy      222 K------VIKPGFIHSPWKSAY------IRQ--HRIDTNWRRGELKSPKV-LKGHDDHVI 266
                     ||    :|  | |      | |    |::||| |     ::  :      :
Db      176 GWGQYLFKNKPPDGKTPPNSFYRALYPKIIQDIETIESNWRCGRHSLQRIHCRSETSKGV 235

Qy      267 TCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLK 326
             |||:   :||||  |||:|:|   | :| | |:||||  |   |  :  :||:||:| |::
Db      236 YCLQYDDQKIVSGLRDNTIKIWDKNTLECKRVLMGHTGSVLCLQYDERVIITGSSDSTVR 295

Qy      327 VWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL---HVLMGHV 383
            ||:  |||  ::||  |   |  :  :   :|: |:| :: |||: :   :    ||:||
Db      296 VWDVNTGEMLNTLIHHCEAVLHLRFNNGMMVTCSKDRSIAVWDMASATDITLRRVLVGHR 355

Qy      384 AAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDT 443
            |||  | :| : :|| : |  :|||:  |   : || ||   :  ||:    ||||| |
Db      356 AAVNVVDFDDKYIVSASGDRTIKVWNTSTCEFVRTLNGHKRGIACLQYRDRLVVSGSSDN 415

Qy      444 SIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTG---------Q 494
            :||:||:| | |:  | ||: |   :   :  :|||  |  :|:||:
Db      416 TIRLWDIECGACLRVLEGHEELVRCIRFDNKRIVSGAYDGKIKVWDLVAALDPRAPAGTL 475

Qy      495 CLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWD 532
            ||:||    :|   |  |||::  :::|| | |: :||
Db      476 CLRTLV---EHSGRVFRLQFDEFQIVSSSHDDTILIWD 510

RESULT  11
POP1_SCHPO
ID   POP1_SCHPO     STANDARD;      PRT;   775 AA.
AC   P87060;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WD-REPEAT PROTEIN POP1.
GN   POP1 OR SPBC2G2.18.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=97347242; PubMed=9203581;
RA   Kominami K., Toda T.;
RT   "Fission yeast WD-repeat protein pop1 regulates genome ploidy through
RT   ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rum1
RT   and the S-phase initiator Cdc18.";
RL   Genes Dev. 11:1548-1560(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Wood V., Rajandream M.A., Barrell B.G., Devlin K., Churcher C.M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN MAINTENANCE OF PLOIDY THROUGH PROTEASOME
CC       DEPENDENT DEGRADATION OF CDK INHIBITOR RUM1 AND S-PHASE INITIATOR
CC       CDC18. FUNCTIONS AS A RECOGNITION FACTOR FOR RUM1 AND CDC18, WHICH
CC       ARE SUBSEQUENTLY UBIQUITINATED AND TARGETED TO THE 26S PROTEASOME
CC       FOR DEGRADATION.
CC   -!- SUBUNIT: BINDS TO CDC18.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y08391; CAA69671.1; -.
DR   EMBL; AL022103; CAA17898.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 3.
DR   PROSITE; PS50082; WD_REPEATS_2; 4.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   DOMAIN      298    345       F-BOX.
FT   REPEAT      444    472       WD 1.
FT   REPEAT      484    538       WD 2.
FT   REPEAT      575    603       WD 3.
FT   REPEAT      615    645       WD 4.
FT   REPEAT      657    687       WD 5.
SQ   SEQUENCE   775 AA;  87816 MW;  B06EDBA46553EEC1 CRC64;

  Query Match             19.7%;  Score 620;  DB 1;  Length 775;
  Best Local Similarity   27.5%;  Pred. No. 5.2e-41;
  Matches  153;  Conservative  100;  Mismatches  179;  Indels  124;  Gaps   16;

Qy      109 SVQPPTGLQEW----LKMFQSWSGPEKL------LALDELIDSCEPTQVKHMMQVIEPQF 158
            |: |   |       ||: |  |  |||        |  |:  |    |::: ::: | |
Db      237 SIHPSNHLSSQKNAVLKLAQLISSFEKLPESVRQYLLFHLLSRCGKHAVQNIHKILLPIF 296

Qy      159 QRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYW-RILAEDNLLW-------------- 203
            |::|::  | |:   ||: |:   |   :|   :| :::: :  ||
Db      297 QKNFLTGFPAEITNLVLTHLDAPSLCAVSQVSHHWYKLVSSNEELWKSLFLKDGFFWDSI 356

Qy      204 ----REKCKEEGIDEPLHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLK 259
                |  | |: :     :||                | |   :   |    | :|: :|
Db      357 DSKIRTMCLEQSLSACAIMKR---------------VYFRHFNLRERW----LHAPEKIK 397

Qy      260 G-----HDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDN 314
                  |   :|| |||  ::|:  :    : ::   ||  :|:|  | | ||: :   :
Db      398 RCSFPIHGVRLITKLQFDDDKIIVSTCSPRINIYDTKTGVLIRSLEEHEGDVWTFEYVGD 457

Qy      315 IIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHL--------------HEKR----- 355
             :::||||||::||:  ||||    ||||||:||: :               | |
Db      458 TLVTGSTDRTVRVWDLRTGECKQVFYGHTSTIRCIKIVQGNQSTTDTDDVEKENRPASND 517

Qy      356 -------VVSGSRDATLRVWDIETGQCL----------------------------HVLM 380
                   ::| ||| |:|:| :    ||                              |
Db      518 ANSMPPYIISSSRDCTIRLWSL---PCLDDPPFVNVNENPDQNNDFTSATTNPFYIRTLR 574

Qy      381 GHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGI--HVVS 438
            ||  :|| |   |  :|| :||  ::||   |  ||| |:||  ||||:  :      :|
Db      575 GHTDSVREVACLGDLIVSASYDGTLRVWKASTGVCLHVLRGHVGRVYSVTINPSRQQCIS 634

Qy      439 GSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNA--DSTVKIWDIKTGQCL 496
               |  ||:|::|:|  :  || || :| | :    ||||| :|  |:::::||: || |
Db      635 AGTDAKIRIWNLESGELLQTLHGHSNLVSQVTFNQNILVSASAPPDTSLRVWDLNTGSCR 694

Qy      497 QTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWR 556
              |: |  |       | ::: |:: |   |::|||:::|: :|:|:|        ::|:
Db      695 DILKCPLGH----IFFQHDESKVVSGS-HSTLQLWDIRSGKLVRDLLT-----DLDIIWQ 744

Qy      557 IRASNTKLVCAVGSRN 572
            :  :    | ||   |
Db      745 VAYNENVCVAAVLRNN 760

RESULT  12
CC4_YEAST
ID   CC4_YEAST      STANDARD;      PRT;   779 AA.
AC   P07834;
DT   01-AUG-1988 (Rel. 08, Created)
```

```
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CELL DIVISION CONTROL PROTEIN 4.
GN   CDC4 OR YFL009W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88011240; PubMed=3309335;
RA   Yochem J., Byers B.;
RT   "Structural comparison of the yeast cell division cycle gene CDC4 and
RT   a related pseudogene.";
RL   J. Mol. Biol. 195:233-245(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RX   MEDLINE=95400292; PubMed=7670463;
RA   Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA   Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA   Yamazaki M., Tashiro H., Eki T.;
RT   "Analysis of the nucleotide sequence of chromosome VI from
RT   Saccharomyces cerevisiae.";
RL   Nat. Genet. 10:261-268(1995).
RN   [3]
RP   SEQUENCE OF 1-579 FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Barrell B.G., Churcher C., Rajandream M.A.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC       REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC       POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC       DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
CC       VARIOUS ASPECTS OF SPORULATION. REQUIRED FOR HTA1-HTB1 LOCUS
CC       TRANSCRIPTION ACTIVATION.
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X05625; CAA29113.1; -.
DR   EMBL; D50617; BAA09229.1; -.
DR   EMBL; D31600; BAA06495.1; -.
DR   EMBL; Z46255; CAA86341.1; -.
DR   PIR; A26867; A26867.
DR   SGD; S0001885; CDC4.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 6.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 4.
DR   PROSITE; PS50082; WD_REPEATS_2; 5.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Cell division; Mitosis; Sporulation; Repeat; WD repeat.
FT   DOMAIN       272    319       F-BOX.
FT   REPEAT       380    408       WD 1.
FT   REPEAT       420    449       WD 2.
FT   REPEAT       461    493       WD 3.
FT   REPEAT       528    556       WD 4.
FT   REPEAT       568    598       WD 5.
FT   REPEAT       630    658       WD 6.
FT   REPEAT       669    698       WD 7.
FT   CONFLICT     460    460       K -> E (IN REF. 1).
SQ   SEQUENCE   779 AA;  86089 MW;  0348F2F8FA78F3BC CRC64;
```

```
  Query Match             19.7%;  Score 618.5;  DB 1;  Length 779;
  Best Local Similarity   29.6%;  Pred. No. 6.9e-41;
  Matches  152;  Conservative  105;  Mismatches  176;  Indels   81;  Gaps   16;

Qy    135 LDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYW- 193
          |  |: : : :::  :  :|:   :|| |: || |::| : ::|: :|:: :    : |
Db    247 LFRLVANMDRSELSDLGTLIKDNLKRDLITSLPFEISLKIFNYLQFEDIINSLGVSQNWN 306

Qy    194 RILAEDNLLWREKCKEEGIDEP-----LHIKRRKVIKPGFIHSPWKSAYIRQHRIDTNWR 248
          :|: :   ||::    |    |     |::|  :           : :::    |  ||
Db    307 KIIRKSTSLWKKLLISENFVSPKGFNSLNLKLSQKYPKLSQQDRLRLSFLENIFILKNW- 365

Qy    249 RGELKSPK------VLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGH 302
               :||       |:||   |||||||  | :::|:||  ::|: ::  | |  | ||
Db    366 ----YNPKFVPQRTTLRGHMTSVITCLQFEDNYVITGADDKMIRVYDSINKKFLLQLSGH 421

Qy    303 TGGVWS-SQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHE----KRVV 357
           ||||:       |::|||||||::||: : | | |   || |||||: : |    | :|
Db    422 DGGVWALKYAHGGILVSGSTDRTVRVWDIKKGCCTHVFKGHNSTVRCLDIVEYKNIKYIV 481

Qy    358 SGSRDATLRVWDI-----------------------ETGQCLHVLMGHVAAVRCVQYDGR 394
          :|||| || || :                       |    : || ||:|:|| |   |
Db    482 TGSRDNTLHVWKLPKESSVPDHGEEHDYPLVFHTPEENPYFVGVLRGHMASVRTVSGHGN 541

Qy    395 RVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIH----VVSGSLDTSIRVWDV 450
           ||||:||  : |||     ||: | |||:|:||  :|  |     :| |:||:||:||:
Db    542 IVVSGSYDNTLIVWDVAQMKCLYILSGHTDRIYSTIYD--HERKRCISASMDTTIRIWDL 599

Qy    451 E----TGNC----------------IHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDI 490
          |     | |                ::|| || :|   : | |  |||  || ::: ||
Db    600 ENIWNNGECSYATNSASPCAKILGAMYTLQGHTALVGLLRLSDKFLVSAAADGSIRGWD- 658

Qy    491 KTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGS 550
                :     : : ||:|    : | ::: |::    :::|::|: :   :  ::
Db    659 -ANDYSRKFSYHHTNLSAITTFYVSDNILVSGSEN-QFNIYNLRSGKLVHANILKDA--- 713

Qy    551 GGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDF 584
             :| :      || ||  ::|  :: | :|||
Db    714 -DQIWSVNFKGKTLVAAV-EKDG--QSFLEILDF 743
```

```
RESULT  13
HET1_PODAN
ID   HET1_PODAN     STANDARD;      PRT;  1356 AA.
AC   Q00808;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   VEGETATIBLE INCOMPATIBILITY PROTEIN HET-E-1.
GN   HET-E1.
OS   Podospora anserina.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariales; Sordariaceae; Podospora.
OX   NCBI_TaxID=5145;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96009891; PubMed=7557402;
RA   Saupe S., Turcq B., Begueret J.;
RT   "A gene responsible for vegetative incompatibility in the fungus
RT   Podospora anserina encodes a protein with a GTP-binding motif and G
RT   beta homologous domain.";
RL   Gene 162:135-139(1995).
CC   -!- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH
CC       SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,
CC       HET-C.
CC   -!- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L28125; AAA85775.1; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 10.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00678; WD_REPEATS_1; 10.
DR   PROSITE; PS50082; WD_REPEATS_2; 10.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   GTP-binding; Repeat; WD repeat.
FT   NP_BIND     300    307       GTP (POTENTIAL).
FT   REPEAT      839    869       WD 1.
FT   REPEAT      881    911       WD 2.
FT   REPEAT      923    953       WD 3.
FT   REPEAT      965    995       WD 4.
FT   REPEAT     1007   1037       WD 5.
FT   REPEAT     1049   1079       WD 6.
FT   REPEAT     1091   1121       WD 7.
FT   REPEAT     1133   1163       WD 8.
FT   REPEAT     1175   1205       WD 9.
FT   REPEAT     1217   1247       WD 10.
SQ   SEQUENCE   1356 AA;  149765 MW;  965FB319844E0651 CRC64;

  Query Match            18.5%;  Score 582.5;  DB 1;  Length 1356;
  Best Local Similarity  28.9%;  Pred. No. 1e-37;
  Matches  169;  Conservative   85;  Mismatches  193;  Indels  137;  Gaps   23;

Qy    113 PTGLQEWLKM-------------FQSWSGPEKLL----------ALDELIDSCEPTQVKH 149
          |  |:| :|:             |   |  : ||          |  |  :   || ::
Db    590 PESLEEIVKLCGSFLIIRERTVYFVHQSAKDFLLGTASDKASNKASQEAFELVFPTGIED 649

Qy    150 M--------MQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYW-------- 193
          :        : |:  : :||   |      :  :   :|  |     :| ||
Db    650 VSYIIFWRSLNVMSQKLRRDIYCLNAPGFLIDNVRVPDPDPLATVRYSCIYWIDHLRDLV 709

Qy    194 ------------------RILAEDNLLWRE-----KCKEEGID-----EPL--HIKRRK- 222
                            | |    | | |     :   |||:     | |  |  | :
Db    710 SSTSSKWVHLLQDDGDIHRFLTTKYLYWLEALSLLRALPEGINAIRQLESLLGHTIRGRL 769

Qy    223 --VIKPGF----------------------IHSPWKSAYIRQHR------------IDTN 246
            ::: |:                      : :|  |   :  :            ::
Db    770 IAIVRDGYRFALSYRMIIEKAPLQAYTSALVFAPTDSMIKKIFKKEEPGWISTISVVEAE 829

Qy    247 WRRGELKSPKVLKGHDDHVIT-CLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGG 305
          |        : |:||   |::      | |: ||||| |:|:|   :|   :|| || |
Db    830 WN----ACTQTLEGHGSSVLSVAFSADGQRVASGSDDKTIKIWDTASGTGTQTLEGHGGS 885

Qy    306 VWS---SQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHE--KRVVSGS 360
          |||   |  |:  : ||| |:|:|:|:| :| |  || ||   |: :      :|| |||
Db    886 VWSVAFSPDRER-VASGSDDKTIKIWDAASGTCTQTLEGHGGRVQSVAFSPDGQRVASGS 944

Qy    361 RDATLRVWDIETGQCLHVLMGHVAAVRCVQY--DGRRVVSGAYDFMVKVWDPETETCLHT 418
           | |:::||  :| |   | || ::|  | :  ||:|| ||: |  :|:||  : ||  |
Db    945 DDHTIKIWDAASGTCTQTLEGHGSSVLSVAFSPDGQRVASGSGDKTIKIWDTASGTCTQT 1004

Qy    419 LQGHTNRVYSLQF--DGIHVVSGSLDTSIRVWDVETGNCIHTLTGH----QSLTSGMELK 472
          |:||   |:|: |  ||  | ||| | :|::||  :| |  || ||    ||:    :
Db   1005 LEGHGGSVWSVAFSPDGQRVASGSDDKTIKIWDTASGTCTQTLEGHGGWVQSVVFSPD-- 1062

Qy    473 DNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVKL 530
             : ||: | |:||||  :| | |||:|   |  :|  : |: :   | : | |||:|:
Db   1063 GQRVASGSDDHTIKIWDAVSGTCTQTLEG---HGDSVWSVAFSPDGQRVASGSIDGTIKI 1119

Qy    531 WDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGT 574
          ||  :|   :   |||  | || |  :  |      | || :||
Db   1120 WDAASGTCTQ---TLE--GHGGWVHSVAFSPDGQRVASGSIDGT 1158


RESULT 14
SCO2_NEUCR
ID   SCO2_NEUCR     STANDARD;      -  PRT;   650 AA.
AC   Q01277;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   SULFUR CONTROLLER-2 (SCON2).
GN   SCON-2.
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=74-OR23-1A;
RX   MEDLINE=95241499; PubMed=7724564;
RA   Kumar A., Paietta J.V.;
RT   "The sulfur controller-2 negative regulatory gene of Neurospora
RT   crassa encodes a protein with beta-transducin repeats.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).
CC   -!- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
CC   -!- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
CC   -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC   -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U17251; AAA68968.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00646; F-box; 1.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 2.
DR   PROSITE; PS50082; WD_REPEATS_2; 6.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Transcription regulation; Repeat; WD repeat.
FT   DOMAIN      124    170       F-BOX.
FT   REPEAT      292    320       WD 1.
FT   REPEAT      332    360       WD 2.
FT   REPEAT      372    400       WD 3.
FT   REPEAT      411    441       WD 4.
FT   REPEAT      453    488       WD 5.
FT   REPEAT      528    564       WD 6.
FT   REPEAT      576    604       WD 7.
FT   REPEAT      616    644       WD 8.
SQ   SEQUENCE   650 AA;  72189 MW;  7473859C99F1B028 CRC64;

  Query Match            18.3%;  Score 576.5;  DB 1;  Length 650;
  Best Local Similarity  26.3%;  Pred. No. 1.1e-37;
  Matches  155;  Conservative   84;  Mismatches  177;  Indels  173;  Gaps   17;

Qy    106 RITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDS-CEPTQVKHMMQVIEPQFQRDFIS 164
          ::||         |  :| :     : | |  ::   | | |:  : : :    : ||||
Db     71 KLTSADQQAVTHVW-SLFSAAPARHRDLMLQGILSQLCFP-QLSFVSREVNEALKIDFIS 128

Qy    165 LLPKELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLW--------REKCKEEGIDEPL 216
           || |||  || :|:   | :|||  : || ||: : :|          || : |   ||
Db    129 ALPVELAQKVLCYLDTVSLTKAAQVSQRWRTLADSDAVWVRMCEQHVNRKCTKCGWGLPL 188

Qy    217 --------HIKRRKVIKPG----------------------------------------- 227
                  : ::|:: | |
Db    189 LERKKLRNYTRQRQLAKGGPQGRVTELADSHDSQDRSVNQHGKRPAAEAEEEDPIKKRQC 248
```

```
Qy      228 --------FIHSP----WKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNR 275
                     :  |    ||: |  : ::  ||:    |   |||||:: | ||||   |
Db      249 MAAAEASKAVTQPKTRSWKAVYRDRWQVSYNWKNSRYKL-SVLKGHENGV-TCLQLDDNI 306

Qy      276 IVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGEC 335
            : :|| | |:|:|:  | :|:||||||| |: : |  |: :|||| | |:||||  ||||
Db      307 LATGSYDTTIKIWNIETEECIRTLVGHTAGIRALQFDDSKLISGSLDHTIKVWNWHTGEC 366

Qy      336 IHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYD--G 393
            : |   || :|  :|     : ||| | |::::|  : :  : | ||   |     |
Db      367 LSTFAAHTDSVISVHFDGHLLASGSSDKTVKIFDFNSKE-TYCLKGHSDWVNSTHVDIKS 425

Qy      394 RRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---------------------- 431
            | | | : |  :|:|| :|   : | :||   |  :
Db      426 RTVFSASDDTTIKLWDLDTRQVIRTYEGHVGHVQQVLILPPEYEPDEEVLNGASQDNQDA 485

Qy      432 -----------------------DGIH--------------------------------- 435
                                    | |
Db      486 MSVSSGGSGSPSMSHAQIERAGSPGSHSSSHNLLPSSLPSGDEDVRHLYGSAFVADESRP 545

Qy      436 -----VVSGSLDTSIRVWDVETGNCIHTLTGH----QSLTSGMELKDNI-LVSGNADSTV 485
                  ::| ||:::|:||  || |: || ||     ||       | | ::||  |  |
Db      546 LPPRYFMTGGLDSTMRLWDSATGRCLRTLFGHLEGVWSLAG-----DTIRVISGANDGMV 600

Qy      486 KIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLK 534
            | |: ::|:|  |  |   |   |||:  : : : : |:|||::|   |
Db      601 KTWEPRSGKCDATYTG---HCGPVTCVGLSDSLMASGSEDGTIRLHSFK 646


RESULT  15
KMHB_DICDI
ID   KMHB_DICDI     STANDARD;      PRT;   732 AA.
AC   P90648;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   MYOSIN HEAVY CHAIN KINASE B (EC 2.7.1.129) (MHCK B).
GN   MHKB OR MHCKB.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RX   MEDLINE=97277316; PubMed=9115238;
RA   Clancy C.E., Mendoza M.G., Naismith T.V., Kolman M.F., Egelhoff T.T.;
RT   "Identification of a protein kinase from Dictyostelium with homology
RT   to the novel catalytic domain of myosin heavy chain kinase A.";
RL   J. Biol. Chem. 272:11812-11815(1997).
CC   -!- FUNCTION: PHOSPHORYLATES THREONINE IN THE C-TERMINAL TAIL REGION
CC       OF MYOSIN II HEAVY CHAIN. THIS PHOSPHORYLATION IS CRITICAL IN
CC       REGULATING THE ASSEMBLY AND DISASSEMBLY OF MYOSIN II FILAMENT.
CC   -!- CATALYTIC ACTIVITY: ATP + [MYOSIN HEAVY-CHAIN] = ADP +
CC       [MYOSIN HEAVY-CHAIN] PHOSPHATE.
CC   -!- DOMAIN: CONSISTS OF AN N-TERMINAL DOMAIN, A CENTRAL NONREPETITIVE
CC       CATALYTIC DOMAIN, AND A C-TERMINAL DOMAIN WITH SEVEN WD REPEATS.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   -!- SIMILARITY: BELONGS TO THE MHCK / EF-2 PROTEIN KINASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U90946; AAB50136.1; -.
DR   DictyDb; DD01087; mhkB.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00678; WD_REPEATS_1; 5.
```

```
DR   PROSITE; PS50082; WD_REPEATS_2; 6.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW   WD repeat.
FT   NP_BIND     298    303       ATP (POTENTIAL).
FT   DOMAIN      355    381       POLY-ASN.
FT   REPEAT      458    486       WD 1.
FT   REPEAT      500    528       WD 2.
FT   REPEAT      540    568       WD 3.
FT   REPEAT      580    608       WD 4.
FT   REPEAT      620    648       WD 5.
FT   REPEAT      660    688       WD 6.
FT   REPEAT      700    730       WD 7.
SQ   SEQUENCE   732 AA;  83167 MW;  A7233C4BD56D4088 CRC64;

  Query Match             16.4%;  Score 516;  DB 1;  Length 732;
  Best Local Similarity   34.0%;  Pred. No. 7.2e-33;
  Matches  113;  Conservative   72;  Mismatches  111;  Indels   36;  Gaps    7;

Qy      231 SPWKSAYIRQ--HRIDTNWRRGELKS-----PKVL--------------------KGHDD 263
            || :  ::    : ::||  | : ||     |::|                    ||:
Db      403 SPSRQLFVSNDGNTLNTNKERSKSKSIDLEKPEILINNKKKESINLETIKLIETIKGY-- 460

Qy      264 HVITCLQFCGNRIVSGSDDNTLKV--WSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGST 321
            || : |  | | : :|  ||:::|  : :   :|::|| || | | |    |  : |||:
Db      461 HVTSHLCICDNLLFTGCSDNSIRVYDYKSQNMECVQTLKGHEGPVESICYNDQYLFSGSS 520

Qy      322 DRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMG 381
            | ::|||: :   || || ||   |  : |::| : ||| | |::|||::| :| : |
Db      521 DHSIKVWDLKKLRCIFTLEGHDKPVHTVLLNDKYLFSGSSDKTIKVWDLKTLECKYTLES 580

Qy      382 HVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSL 441
            |  ||: :   |: : ||: |  :|||| :|  | :||:|||  | ::   | :: |||
Db      581 HARAVKTLCISGQYLFSGSNDKTIKVWDLKTFRCNYTLKGHTKWVTTICILGTNLYSGSY 640

Qy      442 DTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQG 501
            | :||||::::  |  || ||      | : | :| : : |:|:||||::| :|  ||:|
Db      641 DKTIRVWNLKSLECSATLRGHDRWVEHMVICDKLLFTASDDNTIKIWDLETLRCNTTLEG 700

Qy      502 PNKHQSAVTCLQF--NKNFVITSSDDGTVKLW 531
               | : | ||    :|  ||: | | ::::|
Db      701 ---HNATVQCLAVWEDKKCVISCSHDQSIRVW 729


Search completed: September 25, 2001, 14:45:03
Job time: 208 sec
```

```
                    GenCore version 4.5
           Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          September 25, 2001, 14:41:10 ; Search time 22.58 Seconds
                                                (without alignments)
                                                3451.181 Million cell updates/sec

Title:           US-09-328-877A-8
Perfect score:   3143
Sequence:        1 MSKPGKPTLNHGLVPVDLKS..........SRNGTEETKLLVLDFDVDMK 589

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
Result             Query
  No.    Score     Match Length DB  ID                  Description
-------------------------------------------------------------------------------
   1      2889     91.9    553   4  Q9NUX6              Q9nux6 homo sapien
   2    2062.5     65.6   1326   5  Q9VZF4              Q9vzf4 drosophila
   3      1238     39.4    587   5  O44083              O44083 caenorhabdi
   4       650     20.7    665   5  Q9GNN6              Q9gnn6 caenorhabdi
   5       635     20.2    569  11   Q9Z159              Q9z159 mus musculu
   6       635     20.2    569  11   Q9QUI5              Q9qui5 mus musculu
   7       630     20.0    569  11   Q9R1G7              Q9r1g7 mus musculu
   8       617     19.6    510   5  O44382              O44382 drosophila
   9       617     19.6    510   5  Q9VDE3              Q9vde3 drosophila
  10       533     17.0    506   3  Q9P7V1              Q9p7v1 schizosacch
  11     463.5     14.7    942   5  O96611              O96611 dictyosteli
  12       461     14.7    334   4  Q9NWX7              Q9nwx7 homo sapien
  13       439     14.0    410  13   Q9PTR5              Q9ptr5 gallus gall
  14       436     13.9    410   6  Q9GL51              Q9gl51 sus scrofa
  15       426     13.6    330   4  Q9NUL4              Q9nul4 homo sapien
  16     410.5     13.1    317  10   Q9M2Z2              Q9m2z2 arabidopsis
  17     392.5     12.5    333  10   Q9SY00              Q9sy00 arabidopsis
  18     392.5     12.5    454   4  Q9HA09              Q9ha09 homo sapien
  19     391.5     12.5    454  11   Q9QUH1              Q9quh1 mus musculu
  20       389     12.4    391  11  O35592              O35592 mus musculu
  21       388     12.3    411   5  O96698              O96698 drosophila
  22     386.5     12.3    277  11   Q9R2A6              Q9r2a6 mus musculu
  23     378.5     12.0    502   3  O74855              O74855 schizosacch
  24     366.5     11.7    594   4  Q9H073              Q9h073 homo sapien
  25       366     11.6    515  10   Q9FGS2              Q9fgs2 arabidopsis
  26     365.5     11.6    404   5  Q9NDC9              Q9ndc9 caenorhabdi
  27     349.5     11.1    553   5  Q9VVI0              Q9vvi0 drosophila
  28       348     11.1    579   5  O76734              O76734 dictyosteli
  29       347     11.0    514  11   Q9WUC8              Q9wuc8 rattus norv
  30       347     11.0    690   5  O61585              O61585 strongyloce
  31     342.5     10.9    514   4  O43660              O43660 homo sapien
  32       338     10.8    513  11   O55039              O55039 mus musculu
  33       336     10.7    307   5  Q9V878              Q9v878 drosophila
  34       336     10.7    777   3  Q9USN3              Q9usn3 schizosacch
  35     333.5     10.6    494   5  Q18295              Q18295 caenorhabdi
  36       332     10.6    476  13   O93531              O93531 xenopus lae
  37       331     10.5    436   5  Q9VAK0              Q9vak0 drosophila
  38       329     10.5    514  11   Q9EQD4              Q9eqd4 mus musculu
  39       328     10.4    347   5  Q9VPL0              Q9vpl0 drosophila
  40       328     10.4    482   5  Q9VYQ9              Q9vyq9 drosophila
  41       328     10.4    522   4  O43445              O43445 homo sapien
  42     324.5     10.3    651   3  Q9P7I3              Q9p7i3 schizosacch
  43       324     10.3    514   4  Q9H9A1              Q9h9a1 homo sapien
  44       323     10.3    485   4  Q9NVX2              Q9nvx2 homo sapien
  45       323     10.3    520   4  O43864              O43864 homo sapien


                              ALIGNMENTS


RESULT   1
Q9NUX6
ID   Q9NUX6     PRELIMINARY;      PRT;   553 AA.
AC   Q9NUX6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CDNA FLJ11071 FIS, CLONE PLACE1004937, MODERATELY SIMILAR TO SEL-10
DE   PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA   Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA   Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK001933; BAA91986.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
DR   SMART; SM00256; FBOX; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   553 AA;  62280 MW;  CA829C221986A3F2 CRC64;

  Query Match            91.9%;  Score 2889;  DB 4;  Length 553;
  Best Local Similarity  99.8%;  Pred. No. 2.8e-237;
  Matches  540;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       49 RMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRIT 108
            :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       13 KMKRKLDHGSEVRSFSLGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRIT 72
```

```
Qy       109 SVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPK 168
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        73 SVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPK 132

Qy       169 ELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGF 228
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       133 ELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGF 192

Qy       229 IHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVW 288
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       193 IHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVW 252

Qy       289 SAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRC 348
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       253 SAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRC 312

Qy       349 MHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVW 408
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       313 MHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVW 372

Qy       409 DPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSG 468
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       373 DPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSG 432

Qy       469 MELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTV 528
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       433 MELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTV 492

Qy       529 KLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDM 588
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       493 KLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDM 552

Qy       589 K 589
             |
Db       553 K 553

RESULT   2
Q9VZF4
ID   Q9VZF4      PRELIMINARY;      PRT;  1326 AA.
AC   Q9VZF4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CG15010 PROTEIN.
GN   CG15010.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003480; AAG22246.1; -.
DR   FlyBase; FBgn0035516; CG15010.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; 5.
KW   Repeat; WD repeat.
SQ   SEQUENCE   1326 AA;  141360 MW;  3F42C873CFA3027F CRC64;


  Query Match             65.6%;  Score 2062.5;  DB 5;  Length 1326;
  Best Local Similarity   66.8%;  Pred. No. 1.8e-166;
  Matches  405;  Conservative   54;  Mismatches   98;  Indels   49;  Gaps   10;

Qy        12 GLVPVDLKSAKEPLPHQTVMKI----------FSISIIAQG---------LPFCRRRMKR 52
             |     | :    : ||   |           :|::| : |            ||:|:
Db       733 GFTEYSLTNPSSLMSHQRKRKFNEGRLLDGGDYSVTISSSGEVGGPGSGVSDNCRKRI-- 790

Qy        53 KLDHGSEVRSFS-LGKKPCKVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSV- 110
               |  |  ||   ||  |  |   | :: |   |:||          |    ||   ||
Db       791 AYDFASTPRSSQHLG--PTAVLSVTPSSHLT--SSTP----------GSALGRRTPRSVP 836

Qy       111 ---QPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLP 167
                 ||  || ||  || ||  |:||||| ||| |:|:||:|||:|||||||||||||||
Db       837 SRDNPPPELQHWLAQFQRWSHVERLLALDRLIDHCDPSQVRHMMKVIEPQFQRDFISLLP 896

Qy       168 KELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEGI------DEPLHIKRR 221
             :||||:|||:|||||||:|||||| || | :|||||:|||::  |      | |   ||
Db       897 RELALFVLSYLEPKDLLRAAQTCRSWRFLCDDNLLWKEKCRKAQILAEPRSDRP---KRG 953

Qy       222 KVIKPGFIHSPWKSAYIRQHRIDTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSD 281
             :      | ||||:||:||| |: |||   :: ||||||||||||||||| |||||||||
Db       954 RDGNMPPIASPWKAAYMRQHIIEMNWRSRPVRKPKVLKGHDDHVITCLQFSGNRIVSGSD 1013

Qy       282 DNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYG 341
             |||||||||| |||||||||||||||||||  |||||||||||||||: ::| |:||| |
Db      1014 DNTLKVWSAVNGKCLRTLVGHTGGVWSSQMSGNIIISGSTDRTLKVWDMDSGACVHTLQG 1073

Qy       342 HTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGAY 401
             |||||||||||  :|||||||||||||||| | |||||:||:||||||||||: :|||||
Db      1074 HTSTVRCMHLHGSKVVSGSRDATLRVWDIEQGSCLHVLVGHLAAVRCVQYDGKLIVSGAY 1133

Qy       402 DFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTG 461
             |:|||:| || : |||||||||||||||||||:||||||||||||||||||||| ||| |
Db      1134 DYMVKIWHPERQECLHTLQGHTNRVYSLQFDGLHVVSGSLDTSIRVWDVETGNCKHTLMG 1193

Qy       462 HQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVIT 521
             ||||||||||: |||||||||||||:||| |||||||| ||||| |||||||||  ||:|
Db      1194 HQSLTSGMELRQNILVSGNADSTVKVWDITTGQCLQTLSGPNKHHSAVTCLQFNSRFVVT 1253
```

```
Qy       522 SSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLV 581
             |||||||||||:|||:|||||| |:|||||||||||||::|||:||||||||||||||:|
Db      1254 SSDDGTVKLWDVKTGDFIRNLVALDSGGSGGVVWRIRANDTKLICAVGSRNGTEETKLMV 1313

Qy       582 LDFDVD 587
             |||||:
Db      1314 LDFDVE 1319


RESULT   3
O44083
ID   O44083      PRELIMINARY;      PRT;   587 AA.
AC   O44083;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   SEL-10.
GN   SEL-10.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98051191; PubMed=9389650;
RA   Hubbard E.J.A., Wu G., Kitajewski J., Greenwald I.;
RT   "sel-10, a negative regulator of lin-12 activity in Caenorhabditis
RT   elegans, encodes a member of the CDC4 family of proteins.";
RL   Genes Dev. 11:3182-3193(1997).
DR   EMBL; AF020788; AAC47809.1; -.
DR   InterPro; IPR000412; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00890; ABC2_MEMBRANE; UNKNOWN_1.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
KW   Repeat; WD repeat.
SQ   SEQUENCE   587 AA;  65311 MW;  2D3970B4EFAA1B8C CRC64;

  Query Match             39.4%;  Score 1238;  DB 5;  Length 587;
  Best Local Similarity   45.6%;  Pred. No. 7.5e-97;
  Matches  262;  Conservative   85;  Mismatches  162;  Indels   66;  Gaps   11;

Qy        55 DHGSEVRSFSLGKKPC----KVSEYTSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSV 110
             |:| |  |:| |        |:|        :  ||:|:   ||
Db        31 DNGEE-SSYSNGSSSSYNADKLSSSRPLQHKLDLSASPSRNNDLNPR------------- 76

Qy       111 QPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKEL 170
                  ::  : :|:  |  |:: |   |:     | :: :  :||| |||||:| || ||
Db        77 -----VEHLIALFKDLSSAEQMDAFTRLLQESNMTNIRQLRAIIEPHFQRDFLSCLPVEL 131

Qy       171 ALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEE-------------------- 210
              : :|  |   |||: ||  : |::::| : :|:    ||
Db       132 GMKILHNLTGYDLLKVAQVSKNWKLISEIDKIWKSLGVEEFKHHPDPTDRVTGAWQGTAI 191

Qy       211 --GIDEPLHI--------KRRKVIKPG--FIHSPWKSAYIRQHRIDTNWRRGELKSPKVL 258
               |:  | ||        :  |: | |  |  :  || |:|  :|: ||    :    ||
Db       192 AAGVTIPDHIQPCDLNVHRFLKLQKFGDIFERAADKSRYLRADKIEKNWNANPIMGSAVL 251

Qy       259 KGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRD--NII 316
             :||:||||||:|   :  :|:||||||||||    |: : ||||||||||:||:      |
Db       252 RGHEDHVITCMQIHDDVLVTGSDDNTLKVWCIDKGEVMYTLVGHTGGVWTSQISQCGRYI 311

Qy       317 ISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL 376
             :|||||||:|||:   |  :||| ||||||||| :    :|:|||| ||||||:|:|: |
Db       312 VSGSTDRTVKVWSTVDGSLLHTLQGHTSTVRCMAMAGSILVTGSRDTTLRVWDVESGRHL 371

Qy       377 HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIH- 435
               | || |||||||:||  |||| ||| ||:|:  |  |: || || |||||| |:
Db       372 ATLHGHHAAVRCVQFDGTTVVSGGYDFTVKIWNAHTGRCIRTLTGHNNRVYSLLFESERS 431

Qy       436 -VVSGSLDTSIRVWDV---ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIK 491
              | ||||||||||||    |   |:  | ||  ||||||:|: ||||| |||| |::|||
Db       432 IVCSGSLDTSIRVWDFTRPEGQECVALLQGHTSLTSGMQLRGNILVSCNADSHVRVWDIH 491

Qy       492 TGQCLQTLQGPNKHQSAVTCLQ-FNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGS 550
              | |:  | |   |:||:| || | :| | ||||||||||||:: |  ||:||||:|||:
Db       492 EGTCVHMLSG---HRSAITSLQWFGRNMVATSSDDGTVKLWDIERGALIRDLVTLDSGGN 548

Qy       551 GGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFD 585
             || :||: :::| | ||||||| |||||:::||||
Db       549 GGCIWRLCSTSTMLACAVGSRNNTEETKVILLDFD 583


RESULT   4
Q9GNN6
ID   Q9GNN6      PRELIMINARY;      PRT;   665 AA.
AC   Q9GNN6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   LIN-23.
GN   LIN-23.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=11060233;
RA   Kipreos E.T., Gohel S.P., Hedgecock E.M.;
RT   "The Caenorhabditis elegans F-box/WD-repeat protein LIN-23 functions
RT   to limit cell division during development.";
RL   Development 127:5071-5082(2000).
DR   EMBL; AF275253; AAG28037.1; -.
SQ   SEQUENCE   665 AA;  75916 MW;  BF3E9AF51F12ECCC CRC64;

  Query Match             20.7%;  Score 650;  DB 5;  Length 665;
  Best Local Similarity   32.7%;  Pred. No. 8.9e-47;
  Matches  165;  Conservative   87;  Mismatches  186;  Indels   66;  Gaps   15;

Qy        85 SATPTTFGDLRAANGQGQQRR--RITSVQPPTGL-QEWLKMFQSWSGPEKLLALDELIDS 141
             ::|     ||    |:  : :   |  :|   || :| ||    || |:|   :|:::
Db         7 ASTTQQLADLSLTEGEHDEGKPLSIDYLQGHEGLIEEVLK----WSEHEQLDFMDKIVHR 62

Qy       142 CEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQ------TC----- 190
                 |:  :   | |  |||||| ||  |   :| |    | |:: :       |
Db        63 LSHYQLGKVDNFIRPMLQRDFISNLPAHLVELIL-FNVNSDSLKSCEEVSTSWRCALARG 121

Qy       191 RYWRILAEDNL----LW---------------------REKCKEEGIDEPLHIKRRKVIK 225
             ::|: | | |:    ||                     |  |::   |  ::||| |: :
Db       122 QHWKKLIEKNVRSDSLWWGLSEKRQWDKFLNISRDMSVRRICEKFNYD--VNIKRDKLDQ 179

Qy       226 PGFIHSPWKSAY---IRQ-HRIDTNWRRGELKSPKV-LKGHDDHVITCLQFCGNRIVSGS 280
                :|  :   |   ||  | || ||:||  |  ::  :  :   : |||:  ::||||
Db       180 LILMHVFYSKLYPKIIRDIHNIDNNWKRGNYKMTRINCQSENSKGVYCLQYDDDKIVSGL 239

Qy       281 DDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLY 340
              |||:|:|      | | | |||| |   |  : :|||||:| |::||: |||||| ||
Db       240 RDNTIKIWDRKDYSCSRILSGHTGSVLCLQYDNRVIISGSSDATVRVWDVETGECIKTLI 299

Qy       341 GHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL---HVLMGHVAAVRCVQYDGRRVV 397
              |   |  :      :|: |:| :: |||: : : :    ||:|| |||  | :| | :|
Db       300 HHCEAVLHLRFANGIMVTCSKDRSIAVWDMVSPRDITIRRVLVGHRAAVNVVDFDDRYIV 359

Qy       398 SGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIH 457
             | : |  :|||  :|   : || ||   :  ||: |  ||||| | :||:||: :| |:
Db       360 SASGDRTIKVWSMDTLEFVRTLAGHRRGIACLQYRGRLVVSGSSDNTIRLWDIHSGVCLR 419

Qy       458 TLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQ---------CLQTLQGPNKHQSA 508
```

```
           | ||: |   :   :  :|||  |  :|:||::            || :|    :|
Db    420 VLEGHEELVRCIRFDEKRIVSGAYDGKIKVWDLQAALDPRALSSEICLCSLV---QHTGR 476

Qy    509 VTCLQFNKNFVITSSDDGTVKLWD 532
          |  |||:   :::|| | |: :||
Db    477 VFRLQFDDFQIVSSSHDDTILIWD 500


RESULT   5
Q9Z159
ID   Q9Z159      PRELIMINARY;      PRT;   569 AA.
AC   Q9Z159;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   BETA-TRANSDUCIN REPEAT CONTAINING PROTEIN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99145465; PubMed=9990853;
RA   Spencer E., Jiang J., Chen Z.J.;
RT   "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
RT   Slimb/beta-TrCP.";
RL   Genes Dev. 13:284-294(1999).
DR   EMBL; AF112979; AAD04181.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_6.
DR   SMART; SM00320; WD40; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   569 AA;  65047 MW;  BC7C7A44815BED96 CRC64;

  Query Match             20.2%;  Score 635;  DB 11;  Length 569;
  Best Local Similarity   30.3%;  Pred. No. 1.3e-45;
  Matches  164;  Conservative   92;  Mismatches  214;  Indels   72;  Gaps   14;

Qy     53 KLDHGSEVRSFSLGKKPCKV--------SEYTSTTGLV----------PCSATPTTFGDL 94
          |  : ||    : |: | |:          | |   |               |
Db     14 KFMNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCINQETVCLTSTAMKTENCVAKA 73

Qy     95 RAANGQG-----QQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKH 149
          : |||       :||:   | :    |   :| |: ||  :::  :: ||      |  |
Db     74 KLANGTSSMIVPKQRKLSASYEKEKEL--CVKYFEQWSESDQVEFVEHLISQMCHYQHGH 131

Qy    150 MMQVIEPQFQRDFISLLP----KELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWRE 205
          :   ::|  |||||: ||      :|  :||:|: | |  |   |: |  :  | :||::
Db    132 INSYLKPMLQRDFITALPARGLDHIAENILSYLDAKSLCAAELVCKEWYRVTSDGMLWKK 191

Qy    206 KCKE--------EGIDEPLHIKRRK------VIKPGFIHSPWKSAY------IRQ--HRI 243
            :          |: |     ||         ||   ::|  | |      | |    |
Db    192 LIERMVRTDSLWRGLAE-----RRGWGQYLFKNKPPDENAPPNSFYRALYPKIIQDIETI 246

Qy    244 DTNWRRGELKSPKV-LKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGH 302
          ::||| |     ::  :      : |||:   :||||  |||:|:|   | :| | | ||
Db    247 ESNWRCGRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKSTLECKRILTGH 306

Qy    303 TGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 362
          || |   |  : :||:||:| |::||:   || ::||  |   |  :  :   :|: |:|
Db    307 TGSVLCLQYGERVIITGSSDSTVRVWDVNAGEMLNTLIHHCEAVLHLRFNNGMMVTCSKD 366

Qy    363 ATLRVWDIETGQCL---HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTL 419
           :: |||: :   :    ||:|| |||  | :| : :|| : |  :|||:  |   : ||
Db    367 RSIAVWDMASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRTIKVWNTSTCEFVRTL 426

Qy    420 QGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 479
```

```
           ||   :  ||:    ||||| | :||:||:| | |:  | ||: |   :   :  :|||
Db    427 NGHKRGIACLQYRDRLVVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDNKRIVSG 486

Qy    480 NADSTVKIWDIKTG---------QCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKL 530
            |  :|:||:             ||:||    :|   |  |||::  :::|| | |: :
Db    487 AYDGKIKVWDLMAALDPRAPAGTLCLRTLV---EHSGRVFRLQFDEFQIVSSSHDDTILI 543

Qy    531 WD 532
          ||
Db    544 WD 545


RESULT   6
Q9QUI5
ID   Q9QUI5      PRELIMINARY;      PRT;   569 AA.
AC   Q9QUI5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   UBIQUITIN LIGASE FWD1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99199275; PubMed=10097128;
RA   Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA   Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,
RA   Nakayama K.-i.;
RT   "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a
RT   ubiquitin ligase Skp1/Cul 1/F-box protein FWD1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99075339; PubMed=9859996;
RA   Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA   Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;
RT   "Identification of the receptor component of the IkappaBalpha-
RT   ubiquitin ligase.";
RL   Nature 396:590-594(1998).
DR   EMBL; AF081887; AAD17755.1; -.
DR   EMBL; AF099932; AAD08701.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_6.
DR   SMART; SM00320; WD40; 1.
KW   Ligase; Repeat; WD repeat.
SQ   SEQUENCE   569 AA;  65105 MW;  BC7D6544815B2296 CRC64;

  Query Match             20.2%;  Score 635;  DB 11;  Length 569;
  Best Local Similarity   30.3%;  Pred. No. 1.3e-45;
  Matches  164;  Conservative   92;  Mismatches  214;  Indels   72;  Gaps   14;

Qy     53 KLDHGSEVRSFSLGKKPCKV--------SEYTSTTGLV----------PCSATPTTFGDL 94
          |  : ||    : |: | |:          | |   |               |
Db     14 KFMNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCINQETVCLTSTAMKTENCVAKA 73

Qy     95 RAANGQG-----QQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKH 149
          : |||       :||:   | :    |   :| |: ||  :::  :: ||      |  |
Db     74 KLANGTSSMIVPKQRKLSASYEKEKEL--CVKYFEQWSESDQVEFVEHLISQMCHYQHGH 131

Qy    150 MMQVIEPQFQRDFISLLP----KELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWRE 205
          :   ::|  |||||: ||      :|  :||:|: | |  |   |: |  :  | :||::
Db    132 INSYLKPMLQRDFITALPARGLDHIAENILSYLDAKSLCAAELVCKEWYRVTSDGMLWKK 191

Qy    206 KCKE--------EGIDEPLHIKRRK------VIKPGFIHSPWKSAY------IRQ--HRI 243
            :          |: |     ||         ||   ::|  | |      | |    |
```

```
Db    192 LIERMVRTDSLWRGLAE-----RRGWGQYLFKNKPPDENAPPNSFYRALYPKIIQDIETI 246

Qy    244 DTNWRRGELKSPKV-LKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGH 302
          ::||| |     ::  :       :|||:   :||||  |||:|:|   | :| | | ||
Db    247 ESNWRCGRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKSTLECKRILTGH 306

Qy    303 TGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 362
          || |   |  : :||:||:| |::||:   || ::||  |   |  :  :   :|: |:|
Db    307 TGSVLCLQYDERVIITGSSDSTVRVWDVNAGEMLNTLIHHCEAVLHLRFNNGMMVTCSKD 366

Qy    363 ATLRVWDIETGQCL---HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTL 419
           :: |||: :   :    ||:|| |||  | :| : :|| : |  :|||:  |   : ||
Db    367 RSIAVWDMASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRTIKVWNTSTCEFVRTL 426

Qy    420 QGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 479
           ||   :  ||:    ||||| | :||:||:| | |:  | ||: |   :   :  :|||
Db    427 NGHKRGIACLQYRDRLVVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDNKRIVSG 486

Qy    480 NADSTVKIWDIKTG---------QCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKL 530
            |  :|:||:             ||:||    :|   |  |||::  :::|| | |: :
Db    487 AYDGKIKVWDLMAALDPRAPAGTLCLRTLV---EHSGRVFRLQFDEFQIVSSSHDDTILI 543

Qy    531 WD 532
          ||
Db    544 WD 545

RESULT   7
Q9R1G7
ID   Q9R1G7      PRELIMINARY;      PRT;   569 AA.
AC   Q9R1G7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Winston J., Elledge S.J., Harper J.W.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF110396; AAD41025.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_6.
DR   SMART; SM00320; WD40; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   569 AA;  65209 MW;  E6DDCAD28D551D9D CRC64;

  Query Match              20.0%;  Score 630;  DB 11;  Length 569;
  Best Local Similarity   30.1%;  Pred. No. 3.6e-45;
  Matches  163;  Conservative   91;  Mismatches  216;  Indels   72;  Gaps   14;

Qy     53 KLDHGSEVRSFSLGKKPCKV--------SEYTSTTGLV----------PCSATPTTFGDL 94
          |  : ||    : |: | |:          | |   |               |
Db     14 KFMNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCINQETVCLTSTAMKTENCVAKA 73

Qy     95 RAANGQG-----QQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKH 149
          : |||       :||:   | :    |   :| |: ||  :::  :: ||      |  |
Db     74 KLANGTSSMIVPKQRKLSASYEKEKEL--CVKYFEQWSESDQVEFVEHLISQMCHYQHGH 131

Qy    150 MMQVIEPQFQRDFISLLP----KELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWRE 205
          :   ::|  |||||: ||      :   :||:|: | |  |   |: |  :  | :||::
Db    132 INSYLKPMLQRDFITALPARGLDHITENILSYLDAKSLCAAELVCKEWYRVTSDGMLWKK 191

Qy    206 KCKE--------EGIDEPLHIKRRK------VIKPGFIHSPWKSAY------IRQ--HRI 243
```

```
            :          |: |     ||         ||   ::|  | |       | |   |
Db    192 LIERMVRTDSLWRGLAE-----RRGWGQYLFKNKPPDENAPPNSFYRALYPKIIQDIETI 246

Qy    244 DTNWRRGELKSPKV-LKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGH 302
          ::||| |     ::  :       :|||:   :||||  |||:|:|   | :| | | ||
Db    247 ESNWRCGRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKSTLECKRILTGH 306

Qy    303 TGGVWSSQMRDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 362
          || |   |  : :||:||:| |::||:   || ::||  |   |  :  :   :|: |:|
Db    307 TGSVLCLQYDERVIITGSSDSTVRVWDVNAGEMLDTLTHHCEAVLHLRFNNGMMVTCSKD 366

Qy    363 ATLRVWDIETGQCL---HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTL 419
           :: |||: :   :    ||:|| |||  | :| : :|| : |  :|||:  |   : ||
Db    367 RSIAVWDMASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRTIKVWNTSTCEFVRTL 426

Qy    420 QGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 479
           ||   :  ||:    ||||| | :||:||:| | |:  | ||: |   :   :  :|||
Db    427 NGHKRGIACLQYRDRLVVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDNKRIVSG 486

Qy    480 NADSTVKIWDIKTG---------QCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKL 530
            |  :|:||:             ||:||    :|   |  |||::  :::|| | |: :
Db    487 RYDGKIKVWDLMAALDPRAPAGTLCLRTLV---EHSGRVFRLQFDEFQIVSSSHDDTILI 543

Qy    531 WD 532
          ||
Db    544 WD 545

RESULT   8
O44382
ID   O44382      PRELIMINARY;      PRT;   510 AA.
AC   O44382;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   SLIMB.
GN   SLMB OR SLIMB OR CG3412.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98121115; PubMed=9461217;
RA   Jiang J., Struhl G.;
RT   "Regulation of the Hedgehog and Wingless signalling pathways by the F-
RT   box/WD40-repeat protein Slimb.";
RL   Nature 391:493-496(1998).
DR   EMBL; AF032878; AAC38852.1; -.
DR   FlyBase; FBgn0023423; slmb.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
DR   SMART; SM00320; WD40; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   510 AA;  58952 MW;  DBB0243D3730A5E8 CRC64;

  Query Match              19.6%;  Score 617;  DB 5;  Length 510;
  Best Local Similarity   30.6%;  Pred. No. 3.9e-44;
  Matches  146;  Conservative   91;  Mismatches  190;  Indels   50;  Gaps   10;

Qy    104 RRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFI 163
          |:: :|    |  :   : |  ||   ::  :: |:      |   :   ::|  |||||
Db     30 RKKDSSPTYQTERELCFQYFTQWSESGQVDFVEHLLSRMCHYQHGQINAYLKPMLQRDFI 89

Qy    164 SLLP----KELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWRE----KCKEEGIDEP 215
          :|||      :   :||:|: : |  :   |: |  :  : :||::    | : : :
Db     90 TLLPIKGLDHIGENILSYLDAESLKSSELVCKEWLRVISEGMLWKKLIERKVRTDSLWRG 149
```

```
Qy     216 LHIKRRKVIKPGFIHSPWKSAYIRQH----------------RIDTNWRRGELKSPKV-LK 259
           |  :||  ::  ::  |      | |                |: ||| |     ::  :
Db     150 L-AERRNWMQ--YLFKPRPGQTQRPHSFHRELFPKIMNDIDSIENNWRTGRHMLRRINCR 206

Qy     260 GHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISG 319
             :   : |||:   :||||  |||:|:|     :|::||:|||| |   |  | :||||
Db     207 SENSKGVYCLQYDDGKIVSGLRDNTIKIWDRTDLQCVKTLMGHTGSVLCLQYDDKVIISG 266

Qy     320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL--- 376
           |:| |::||:  ||| ::||  |   |  :  :   :|: |:| :: |||: :   :
Db     267 SSDSTVRVWDVNTGEMVNTLIHHCEAVLHLRFNNGMMVTCSKDRSIAVWDMTSPSEITLR 326

Qy     377 HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHV 436
            ||:|| |||  | :| : :|| : |  :|||   |   : || ||   :  ||:    |
Db     327 RVLVGHRAAVNVVDFDEKYIVSASGDRTIKVWSTSTCEFVRTLNGHKRGIACLQYRDRLV 386

Qy     437 VSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDI------ 490
           |||| | |||:||:| | |:  | ||: |   :      :|||  |  :|:||:
Db     387 VSGSSDNSIRLWDIECGACLRVLEGHEELVRCIRFDTKRIVSGAYDGKIKVWDLVAALDP 446

Qy     491 ---KTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDL--------KTG 536
                  || ||    :|   |  ||::   :::|| | |: :||         |||
Db     447 RAASNTLCLNTLV---EHTGRVFRLQFDEFQIVSSSHDDTILIWDFLNFTPNENKTG 500


RESULT   9
Q9VDE3
ID   Q9VDE3      PRELIMINARY;      PRT;   510 AA.
AC   Q9VDE3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   SLMB PROTEIN (SLIMB).
GN   SLMB OR CG3412.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
```

```
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY, AND IMAGINAL DISC;
RX   MEDLINE=20245299; PubMed=10781936;
RA   Miletich I., Limbourg-Bouchon B.;
RT   "Drosophila null slimb clones transiently deregulate Hedgehog-
RT   independent transcription of wingless in all limb discs, and induce
RT   decapentaplegic transcription linked to imaginal disc regeneration.";
RL   Mech. Dev. 93:15-26(2000).
DR   EMBL; AE003733; AAF55853.1; -.
DR   EMBL; AF222924; AAF63214.1; -.
DR   EMBL; AF222923; AAF63213.1; -.
DR   FlyBase; FBgn0023423; slmb.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; 5.
DR   SMART; SM00256; FBOX; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   510 AA;  58952 MW;  F4D5DF126F58A012 CRC64;


  Query Match             19.6%;  Score 617;  DB 5;  Length 510;
  Best Local Similarity   30.6%;  Pred. No. 3.9e-44;
  Matches  146;  Conservative   92;  Mismatches  189;  Indels   50;  Gaps   10;

Qy     104 RRRITSVQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFI 163
           |:: :|    |  :   : |  ||   ::  ::|:       |   :   ::|  |||||
Db      30 RKKDSSPTYQTERELCFQYFTQWSESGQVDFVEHLLSRMCHYQHGQINAYLKPMLQRDFI 89

Qy     164 SLLP----KELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWRE----KCKEEGIDEP 215
           :|||      :|  :||:|: : |  :   |: |  :  : :||::    | : : :
Db      90 TLLPIKGLDHIAENILSYLDAESLKSSELVCKEWLRVISEGMLWKKLIERKVRTDSLWRG 149

Qy     216 LHIKRRKVIKPGFIHSPWKSAYIRQH----------------RIDTNWRRGELKSPKV-LK 259
           |  :||  ::  ::  |      | |                |: ||| |     ::  :
Db     150 L-AERRNWMQ--YLFKPRPGQTQRPHSFHRELFPKIMNDIDSIENNWRTGRHMLRRINCR 206

Qy     260 GHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMRDNIIISG 319
             :   : |||:   :||||  |||:|:|     :|::||:|||| |   |  | :||||
Db     207 SENSKGVYCLQYDDGKIVSGLRDNTIKIWDRTDLQCVKTLMGHTGSVLCLQYDDKVIISG 266

Qy     320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL--- 376
           |:| |::||:  ||| ::||  |   |  :  :   :|: |:| :: |||: :   :
Db     267 SSDSTVRVWDVNTGEMVNTLIHHCEAVLHLRFNNGMMVTCSKDRSIAVWDMTSPSEITLR 326

Qy     377 HVLMGHVAAVRCVQYDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQFDGIHV 436
            ||:|| |||  | :| : :|| : |  :|||   :   : || ||   :  ||:    |
Db     327 RVLVGHRAAVNVVDFDEKYIVSASGDRTIKVWSTSSCEFVRTLNGHKRGIACLQYRDRLV 386

Qy     437 VSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDI------ 490
           |||| | |||:||:| | |:  | ||: |   :      :|||  |  :|:||:
Db     387 VSGSSDNSIRLWDIECGACLRVLEGHEELVRCIRFDTKRIVSGAYDGKIKVWDLVAALDP 446

Qy     491 ---KTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDL--------KTG 536
                  || ||    :|   |  ||::   :::|| | |: :||         |||
Db     447 RAASNTLCLNTLV---EHTGRVFRLQFDEFQIVSSSHDDTILIWDFLNFTPNENKTG 500


RESULT  10
Q9P7V1
```

```
ID   Q9P7V1      PRELIMINARY;      PRT;   506 AA.
AC   Q9P7V1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WD-REPEAT PROTEIN.
GN   SPAC30.05.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   McDougall R.C., Rajandream M.A., Barrell B.G., Brown S., Murphy L.,
RA   Jones L., McNeil A., Harris D.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL136538; CAB66464.1; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_3.
DR   SMART; SM00256; FBOX; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   506 AA;  58256 MW;  CEF34D4EFFBC2E10 CRC64;

  Query Match             17.0%;  Score 533;  DB 3;  Length 506;
  Best Local Similarity   25.8%;  Pred. No. 5.4e-37;
  Matches  140;  Conservative   91;  Mismatches  204;  Indels  108;  Gaps   14;

Qy     55 DHGSEVRSFSLGKKPCKVSEYT-----STTGLVPCSATPTTFGDLRAANGQGQQRRRITS 109
          |:|  | |        |||: |     ||:  |||             |    :  ||
Db      5 DNGKNVVS--------KVSDLTSCSDFSTSSPVPC------------LNPLSHENNRIDL 44

Qy    110 VQPPTGLQEWLKMFQSWSGPEKLLALDELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKE 169
          ::                         :|: |     |  :   :      ||  : |:|
Db     45 IR-------------------------DLLASLSKEGVVAVYNHVRSLLFTDFTEVFPEE 79

Qy    170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEEG--IDEPL---------- 216
          ::| | |:|:  || :     : |: | ||  :|:    ::|  ::| :
Db     80 VSLRVFSYLDQLDLCKCKLMSKRWKRLLEDPGIWKALYMQKGWFVNENVLNEFEAWRRTH 139

Qy    217 ---------HIKRRKVIKP-GFIHSP-------------WKSAYIRQHRIDTNWRRGE- 251
                    :|:: :| | | :  |              |   |     :|:||| |
Db    140 KFPQPRFENFLKQQNIIGPYGTMFLPQQFIFDSNGRPLLNWSYLYKEHAHLDSNWRHGRF 199

Qy    252 ----LKSPKVLKGHDDHV------ITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVG 301
                :| :    |         : |:|:    :|||| | |: ||   :   |  | |
Db    200 LVSTFNNPSIRFPADQDFRATLDSVYCVQYDDEIMVSGSKDRTVSVWDVNSRFILYKLYG 259

Qy    302 HTGGVWSSQM--RDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSG 359
          |:| |       | |:::|||:| |: :|: :    :   :|||  |  : : |  ::|
Db    260 HSGSVLCLDFCRRRNLLVSGSSDSTIIIWDWQNRRPLKVYFGHTDNVLGVVVSENYIISS 319

Qy    360 SRDATLRVWDIETGQ----CLHVLMGHVAAVRCVQYDGRR--VVSGAYDFMVKVWDPETE 413
          ||| | ||| ::       |:||| ||:|:|  |||  :   :|: : |  :: ||  |
Db    320 SRDHTARVWRLDATSPAEACMHVLRGHLASVNSVQYSSKTGLIVTASSDRTLRTWDITTG 379

Qy    414 TCLHTLQGHTNRVYSLQFDGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
           |:  :  |   :   |::|  :|||| | :||:::  :|  :  | ||: |   :   |
Db    380 HCIRIIHAHQRGIACAQYNGKFIVSGSSDLTIRIFEASSGKLLRMLQGHEDLIRTVRFND 439

Qy    474 NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSSDDGTVKLWDL 533
            :|||  | ||:||:  ||:    |   |   | |  |||:   :|  :    : :|:
Db    440 EKIVSGGYDGTVRIWNFNTGEQHCVLH--NSRNSRVFGLQFDHRRIIACTHSSEILVWNF 497

Qy    534 KTG 536
            |
```

```
Db    498 DDG 500

RESULT 11
O96611
ID   O96611      PRELIMINARY;      PRT;   942 AA.
AC   O96611;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   MEK KINASE ALPHA.
GN   MKKA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX . NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KAX-3;
RX   MEDLINE=99051319; PubMed=9832508;
RA   Chung C.Y., Reddy T.B.K., Zhou K., Firtel R.A.;
RT   "A novel, putative MEK kinase controls developmental timing and
RT   spatial patterning in Dictyostelium and is regulated by ubiquitin-
RT   mediated protein degradation.";
RL   Genes Dev. 12:3564-3578(1998).
CC   -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR   EMBL; AF093689; AAC97114.1; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001245; -.
DR   InterPro; IPR001680; -.
DR   InterPro; IPR001810; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00400; WD40; 7.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS50181; FBOX; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_4.
DR   SMART; SM00320; WD40; 1.
KW   ATP-binding; Kinase; Repeat; Serine/threonine-protein kinase;
KW   Transferase; Tyrosine-protein kinase; WD repeat.
SQ   SEQUENCE   942 AA;  105796 MW;  C9E4928A8C7C68F7 CRC64;

  Query Match             14.7%;  Score 463.5;  DB 5;  Length 942;
  Best Local Similarity   24.6%;  Pred. No. 1e-30;
  Matches  127;  Conservative  118;  Mismatches  212;  Indels   59;  Gaps   18;

Qy     76 TSTTGLVPCSATPTTFGDLRAANGQGQQRRRITSVQPPTGLQEWLKMFQSWSGPEKLLAL 135
          | :|     :|| ||  :  | |    |:::    ||||  |  : :    |  ::
Db    446 TLSTNTTNTTATTTTNN--ATNSNINQQQQQQQQPPTRTQR-VSISAGSSNNKRYTPP 502

Qy    136 DELIDSCEPTQVKHMMQVIEPQFQRDFISLLPKELALYVLSFLEPKDLLQAAQTCRYWRI 195
               |   : : :   :           :||  | : :    ::| :   :: |::|:
Db    503 ISTSTSSSSSSILNNFSI---------NIILPINLIILIFREIKPNFVNTLSRVCKHWKQ 553

Qy    196 LAEDNLLWREKCKEEGIDEPLHIKRRKVIKPGFIHS-PWKSAYIRQHRIDTNWRRGELKS 254
          : :|: || : | :           | : |  |  |  ||| ||: ::    |   :| :
Db    554 IIDDDELWNKYCSD-----------RLINKSKFEESITWKSNYIKIYKQQKVWFHNKL-N 601

Qy    255 PKVLKGHDDHVITCLQFCGNR--IVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVWSSQMR 312
             |||||  |  |::   ::  ::|| :|  ||||  ::|        |:| | |   :
Db    602 HSTLKGHDKGVF-CVKLIDDQGMVLSGGEDKKLKVWD-ISGNHHHN--HHSGIVGSISKK 657

Qy    313 D--NIIISGSTDRTLKVWNAETGEC--IHTLYGHTSTVRCMHLHEK------RVVSGSRD 362
             ||| : ::: :    |: :     :: | ||:  :: :    :      || : | |
Db    658 SGLNIINNNNSNNSNSSSNSSSSNSRYLFSLKGHSGCIKSVDYQRQSGSDVSRVFTASAD 717

Qy    363 ATLRVWDIETGQCLHVLMGHVAAVRCVQYDG---RRVVSGAYDFMVKVWDPETETCLHTL 419
           | ::: ::| : |     |  || |: | |    : :: : |  :::|| || :|| ||
```

```
Db    718 FTCKIFSLKTKKTLFTYTNHQEAVTCINYLGDVENKCITSSLDKTIQLWDAETGSCLSTL 777

Qy    420 QGHTNRVYSLQFDGI--------H-VVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGME 470
          :|||  :| :: | :        | ||| |:| :  |||  : : : : | |
Db    778 RGHTGGIYCVKTDQVATHGNGYNHLVVSASVDKTSNVWDTRSSSKVRSFTQHTEDVLCCY 837

Qy    471 LKDNILVSGNADSTVKIWDIKTGQCLQT-LQGPNKHQSAVTCLQFNKNFVITSSDDGTVK 529
          : |  :|:|: | |:|:||| ||: : | :    : :: |  :||::: :|:|   | ::
Db    838 VFDQKVVTGSCDGTIKLWDIGTGKTISTFIPSETRQKNYVWTVQFDQSKIISSGKTGIIR 897

Qy    530 LWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLV 565
          :||:      |::     ||    :: :: :| ||:
Db    898 IWDIYNERDSRSI-----GGHHETIFSLQFNNQKLI 928

RESULT  12
Q9NWX7
ID   Q9NWX7      PRELIMINARY;      PRT;   334 AA.
AC   Q9NWX7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CDNA FLJ20545 FIS, CLONE KAT11476.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA   Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA   Nakamura Y., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK000552; BAA91248.1; -.
DR   InterPro; IPR001632; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00319; GPROTEINB.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_4.
KW   Repeat; WD repeat.
SQ   SEQUENCE   334 AA;  36588 MW;  4BF30914A2250286 CRC64;

  Query Match             14.7%;  Score 461;  DB 4;  Length 334;
  Best Local Similarity   33.9%;  Pred. No. 4e-31;
  Matches  102;  Conservative   68;  Mismatches   93;  Indels   38;  Gaps   13;

Qy    258 LKGHDDHVITCLQFC--GNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGV----WSSQM 311
          | ||    :: ::|   |  : | | |  :|:| |  ||  :|: ||  |:    |||
Db     41 LAGH-TKAVSSVKFSPNGEWLASSSADKLIKIWGAYDGKFEKTISGHKLGISDVAWSSD- 98

Qy    312 RDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKR--VVSGSRDATLRVWD 369
            |:::| | |:|||:|:  :|:|: || ||:: | | : : :   :|||| | ::|:||
Db     99 -SNLLVSASDDKTLKIWDVSSGKCLKTLKGHSNYVFCCNFNPQSNLIVSGSFDESVRIWD 157

Qy    370 IETGQCLHVLMGHVAAVRCVQY--DGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVY 427
          ::||:||  |  |   |  | :  ||  :|| :|| : ::||  :  || ||    |
Db    158 VKTGKCLKTLPAHSDPVSAVHFNRDGSLIVSSSYDGLCRIWDTASGQCLKTLIDDDNPPV 217

Qy    428 S-LQF--DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQ----------SLTSGMELKDN 474
          | ::|  :| :::: :|| ::::||   | |: | |||:          |:| |
Db    218 SFVKFSPNGKYILAATLDNTLKLWDYSKGKCLKTYTGHKNEKYCIFANFSVTGG-----K 272

Qy    475 ILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAV--TCLQFNKNFVITSS--DDGTVKL 530
           :|||: |: | ||:::| : :| |||   |   |  |     :| : :::  :| |:||
Db    273 WIVSGSEDNLVYIWNLQTKEIVQKLQG---HTDVVISTACHPTENIIASAALENDKTIKL 329

Qy    531 W 531
          |
Db    330 W 330
```

```
RESULT  13
Q9PTR5
ID   Q9PTR5      PRELIMINARY;      PRT;   410 AA.
AC   Q9PTR5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   LIS1.
GN   LIS1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Reiner O., Shmueli O.;
RT   "Characterization of the chicken homolog of LIS1.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF113946; AAF18938.1; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
DR   SMART; SM00320; WD40; 1.
KW   Repeat; WD repeat.
SQ   SEQUENCE   410 AA;  46664 MW;  FC5848D06E0DCA20 CRC64;

  Query Match             14.0%;  Score 439;  DB 13;  Length 410;
  Best Local Similarity   29.9%;  Pred. No. 4e-29;
  Matches  109;  Conservative   63;  Mismatches  133;  Indels   60;  Gaps   10;

Qy    227 GFIHSPWKSAYIRQHRI---------------DTNWRRGELKSPK-----------VLKGHD 262
          | :   | |    | ::               :    |: : ||           | ||
Db     49 GLLEKKWTSVIRLQKKVMELESKLNEAKEEFTSGGPLGQKRDPKEWIPRPPEKYALSGHR 108

Qy    263 DHVITCL-QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVW--SSQMRDNIIISG 319
            |   :     : :|| |:| |:|||   ||   ||| |||  |   |      :: |
Db    109 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHTGKLLASC 168

Qy    320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377
          | | |:|:|: :  ||| |::||   |  : :  :   :|| ||| |:::|:::|| |:
Db    169 SADMTIKLWDFQGFECIRTMHGHDHNVSSVAIMPNGDHIVSASRDKTIKMWEVQTGYCVK 228

Qy    378 VLMGHVAAVRCVQ--YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---- 431
             ||   || |:   ||  : | : |  |:||   |: |   |: | : |  : :
Db    229 TFTGHREWVRMVRPNQDGTLIASCSNDQTVRVWVVATKECKAELREHEHVVECISWAPES 288

Qy    432 ------------------DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
                             |  ::||| | :|::||: || |: || || :   |:
Db    289 SYSTISEATGSETKKSGKPGPFLLSGSRDKTIKMWDISTGMCLMTLVGHDNWVRGVLFHS 348

Qy    474 --NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVK 529
              ::|   | |:::|| |  :|::||   | |:  || | |:|   :|:| | | |||
Db    349 GGKFILSCADDKTLRVWDFKNKRCMKTL---NAHEHFVTSLDFHKTAPYVVTGSVDQTVK 405

Qy    530 LWDLK 534
          :|: :
Db    406 VWECR 410

RESULT  14
Q9GL51
ID   Q9GL51      PRELIMINARY;      PRT;   410 AA.
AC   Q9GL51;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PLATELET-ACTIVATING FACTOR ACETYLHYDROLASE IB-ALPHA SUBUNIT.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nagasaka T., Boulday G., Coupel S., Coulon F., Tesson L.,
RA   Heslan J.-M., Soulillou J.-P., Charreau B.;
RT   "Cloning of porcine PAF-AH Ib-alpha cDNA and expression in endothelial
RT   cells.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF319658; AAG33867.1; -.
KW   Hydrolase.
SQ   SEQUENCE   410 AA;  46654 MW;  A08DAFCD8B8B2719 CRC64;

  Query Match             13.9%;  Score 436;  DB 6;  Length 410;
  Best Local Similarity   29.9%;  Pred. No. 7.1e-29;
  Matches  109;  Conservative   63;  Mismatches  133;  Indels   60;  Gaps   10;

Qy    227 GFIHSPWKSAYIRQHRI--------------DTNWRRGELKSPK----------VLKGHD 262
          | :   | |    | ::               :    |: : ||           | ||
Db     49 GLLEKKWTSVIRLQKKVMELESKLNEAKEEFTSGGPLGQKRDPKEWIPRPPEKYALSGHR 108

Qy    263 DHVITCL-QFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTGGVW--SSQMRDNIIISG 319
            |   :     : :|| |:| |:|||   ||   ||| |||  |   |      :: |
Db    109 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASC 168

Qy    320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377
          | | |:|:|: :  ||| |::||   |  : :  :   :|| :|| |:::|:::|| |:
Db    169 SADMTIKLWDFQGFECIRTMHGHDHNVSSVAIMPNGDHIVSAARDKTIKMWEVQTGYCVK 228

Qy    378 VLMGHVAAVRCVQ--YDGRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVYSLQF---- 431
             ||   || |:   ||  : | : |  |:||   |: |   |: | : |  : :
Db    229 TFTGHREWVRMVRPNQDGTLIASCSNDQTVRVWVVATKECKAELREHEHVVECISWAPES 288

Qy    432 ------------------DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKD 473
                             |  ::||| | :|::||| || |: || || :   |:
Db    289 SYSSISEATGSETKKSGKPGPFLLSGSRDKTIKMWDVSTGMCLMTLVGHDNWVRGVLFHS 348

Qy    474 --NILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKN--FVITSSDDGTVK 529
              ::|   | |:::|| |  :|::||   | |:  || | |:|   :|:| | | |||
Db    349 GGKFILSCADDKTLRVWDYKNKRCMKTL---NAHEHFVTSLDFHKTAPYVVTGSVDQTVK 405

Qy    530 LWDLK 534
          :|: :
Db    406 VWECR 410

RESULT  15
Q9NUL4
ID   Q9NUL4     PRELIMINARY;     PRT;   330 AA.
AC   Q9NUL4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CDNA FLJ11287 FIS, CLONE PLACE1009596, WEAKLY SIMILAR TO VEGETATIBLE
DE   INCOMPATIBILITY PROTEIN HET-E-1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK002149; BAA92110.1; -.
```

```
DR   InterPro; IPR001632; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00319; GPROTEINB.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00678; WD_REPEATS; UNKNOWN_4.
KW   Repeat; WD repeat.
SQ   SEQUENCE   330 AA;  36329 MW;  F78396D3B5B716CA CRC64;

  Query Match             13.6%;  Score 426;  DB 4;  Length 330;
  Best Local Similarity   31.8%;  Pred. No. 3.7e-28;
  Matches   95;  Conservative   71;  Mismatches   99;  Indels   34;  Gaps   12;

Qy    258 LKGHDDHVITCLQFC--GNRIVSGSDDNTLKVWSAVTGKCLRTLVGH----TGGVWSSQM 311
          | || :  :: ::|   |  : | | |  : :| |  ||  :|| ||    :   |||
Db     37 LVGHTE-AVSSVKFSPNGEWLASSSADRLIIIWGAYDGKYEKTLYGHNLEISDVAWSSD- 94

Qy    312 RDNIIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLH--EKRVVSGSRDATLRVWD 369
            : ::| | |:|||:|:  :|:|: || ||:: | | : :     ::||| | |:::|:
Db     95 -SSRLVSASDDKTLKLWDVRSGKCLKTLKGHSNYVFCCNFNPPSNLIISGSFDETVKIWE 153

Qy    370 IETGQCLHVLMGHVAAVRCVQYD--GRRVVSGAYDFMVKVWDPETETCLHTLQGHTNRVY 427
          ::||:||  |  |   |  | ::  |  :|||:|| : ::||  :  || ||    |
Db    154 VKTGKCLKTLSAHSDPVSAVHFNCSGSLIVSGSYDGLCRIWDAASGQCLKTLVDDDNPPV 213

Qy    428 S-LQF--DGIHVVSGSLDTSIRVWDVETGNCIHTLTGHQ----------SLTSGMELKDN 474
          | ::|  :| :::: :|| ::::||   | |: | || :          |:| |
Db    214 SFVKFSPNGKYILTATLDNTLKLWDYSRGRCLKTYTGQKNEKYCIFANFSVTGG-----K 268

Qy    475 ILVSGNADSTVKIWDIKTGQCLQTLQGPNKHQSAVTCLQFNKNFVITSS--DDGTVKLW 531
           :|||: |: | ||:::| : :| |||      :  |    :| : :::  :| |:|||
Db    269 WIVSGSEDNLVYIWNLQTKEIVQKLQGHTDVVISAAC-HPTENLIASAALENDKTIKLW 326

Search completed: September 25, 2001, 14:44:42
Job time: 212 sec
```